

# 高速铁路网络延误传播分析\*

王璞, 谭淮锐, 郭宝, 周慧

(中南大学交通运输工程学院, 湖南长沙 410075)

**摘要:**针对既有方法研究高速铁路网络中延误传播规律的局限性,基于高速列车实绩运行数据和列车时刻表数据构建站点延误传播贝叶斯网络模型。结合复杂网络中的渗流理论研究延误传播团簇的演变规律,并以最大延误传播团簇中的南京南站和第二大延误传播团簇中的长沙南站为例,分析处于渗流突变状态下的站点延误传播贝叶斯网络。提出基于核心延误传播团簇的站点延误状态预测模型。结果表明:网络中的站点可按延误传播特征分为延误发散站点、延误传递站点和延误消散站点。延误发散站点不仅能直接将延误传播至邻近站点,还能通过延误传播链将延误传播至较远地区的站点。由此网络以部分延误发散站点为中心向延误传递站点和延误消散站点进行“辐射式”延误传播。

**关键词:**延误传播;贝叶斯网络;复杂网络;渗流理论

**中图分类号:**N94 **文献标志码:**A **开放科学(资源服务)标识码(OSID):**

**文章编号:**1001-2486(2023)06-157-08



听语音  
与作者互动  
聊科研

## Analysis of high-speed railway network delay propagation

WANG Pu, TAN Huairui, GUO Bao, ZHOU Hui

(School of Traffic & Transportation Engineering, Central South University, Changsha 410075, China)

**Abstract:** Noted the limitation of existing methods to study the delay propagation mechanism in high-speed railway network, the Bayesian network was used to analyze delay dependencies between stations based on high-speed train operation data and schedule data. The percolation theory of complex network was applied to study the evolution of delay propagation clusters. Nanjingnan Railway Station in the largest delay propagation cluster and Changshanan Railway Station in the second largest delay propagation cluster were taken as examples to analyze the network at percolation threshold. Based on core delay propagation clusters, the station delay state prediction model was established. The results show that stations can be divided into 3 categories, namely, delay generator, delay mediator and delay absorber according to their characteristic in the aspect of delay propagation. Delay generator can not only spread the delay to stations close to them, but also spread the delay to the far away stations through delay propagation chain, which makes the network taking part of the delay generator as the center and spreading delay to the delay mediator and the delay absorber in a radiational delay propagation mode.

**Keywords:** delay propagation; Bayesian network; complex network; percolation theory

高速铁路列车的运行状态可能会受到多种因素(极端天气、设备故障、运营管理不当等)的影响,进而发生列车延误<sup>[1]</sup>。列车延误不仅会打乱列车运行计划,影响高速铁路的运营管理;而且会使乘客的出行受到影响,降低乘客的乘车满意度。对高速铁路列车延误进行分析研究,理解高速铁路网络中的延误传播机理,对于制定更为科学的列车运行计划、提高运营管理水平具有十分重要的理论和实际意义。下面首先介绍列车延误研究领域的相关研究工作。

根据列车延误发生原因的不同,列车延误可以分为初始延误和连带延误<sup>[2]</sup>。列车发生初始

延误的主要因素包括天气因素、人为因素和系统性因素<sup>[3]</sup>。列车发生初始延误后,延误传播效应的存在会导致后续列车发生连带延误<sup>[4]</sup>。Meester等<sup>[5]</sup>建立了列车延误传播模型,并从列车初始延误的概率分布推导出连带延误的概率分布。胡思继等<sup>[6]</sup>研究了区段内列车延误传播过程和传播规律,定义了相关延误指标并推导出理论计算公式。庄河等<sup>[7]</sup>运用广铁集团管辖的6条高速铁路实绩数据建立了不同致因情况下初始延误时长分布模型。Wang等<sup>[8]</sup>识别了列车延误关键车站,并提出时间间隔阈值用于构建列车延误传播链,识别列车延误传播是否发生。柏赞等<sup>[9]</sup>

\* 收稿日期:2021-12-14

基金项目:科技部国家重点研发计划资助项目(2016YFB1200401)

作者简介:王璞(1983—),男,河北石家庄人,教授,博士,博士生导师,E-mail:wangpu@csu.edu.cn

考虑了不同运行线路的列车在运行线的交叉区域存在运行干扰的情况,并提出最优接近速度策略使受到延误传播影响的列车迅速通过运行干扰区域,尽快恢复正常运行。曾壹等<sup>[10]</sup>使用区间锁闭时间模型对列车延误区间增加和传播的范围进行预测,并提出列车节能-延误恢复模型,不仅可以降低列车牵引能耗,还能缓解列车的延误传播。

上述研究针对列车间的延误传播,其延误本因是铁路线路的连通性以及列车共用铁路线路的运行机制<sup>[1]</sup>,所以列车间的延误传播会导致站点间的延误传播。为研究站点间的延误传播关系,Lessan 等<sup>[11]</sup>运用武广高铁实绩运行数据构建了三种贝叶斯网络,发现基于数据驱动和领域专家知识的混合方法构建的贝叶斯网络在预测延误时间方面有更高的精度。Corman 等<sup>[12]</sup>对瑞典铁路一段交通走廊线路中的站点延误传播关系进行了研究,并运用领域专家知识构建了贝叶斯网络。但这些研究都只关注单条铁路线路,而不是铁路网络。相较于某条铁路单线,铁路网络具有更复杂的拓扑结构,所以仅使用领域专家知识难以判别大型铁路网络中站点间的延误传播关系。因此,本文采用数据驱动方法构建高速铁路站点延误传播贝叶斯网络,并运用复杂网络中的渗流理论寻找并分析处于渗流突变状态下的延误传播团簇,以探究高速铁路网络中的延误传播机理。

## 1 站点延误传播贝叶斯网络模型

### 1.1 贝叶斯网络理论

贝叶斯网络是一种对因果关系进行推理的概率图模型<sup>[13]</sup>,其中每个节点代表的随机变量可用于表达任何问题的抽象形式,由父节点指向子节点的有向边代表节点之间的相互依赖关系。若贝叶斯网络中存在  $n$  个节点  $(1, 2, \dots, n)$ ,每个节点对应的随机变量为  $X_1, X_2, \dots, X_n$ ,则其联合概率分布为

$$P(X_1, X_2, \dots, X_n) = \prod_{j=1}^n P(X_j | p(X_j)) \quad (1)$$

式中,  $p(X_j)$  代表  $X_j$  的父节点集合。

### 1.2 贝叶斯网络结构学习

在研究铁路单线中站点间的延误传播关系时,学者大多结合领域专家知识建立贝叶斯网络<sup>[11]</sup>。但是在研究大型高速铁路网络中站点间的延误传播关系时,复杂的网络拓扑结构使上述方法难以判别站点间的延误传播关系。本文通过

采用 NO TEARS 贝叶斯网络结构学习算法将有向无环图的结构学习问题表示为一个纯连续优化问题,相较于传统启发式算法更加精准,可以学习到全局最优的贝叶斯网络<sup>[14]</sup>。贝叶斯网络的结构优化问题在数学上表示为以下形式:

$$\begin{aligned} \min_{W \in \mathbb{R}^{d \times d}} F(W) \\ \text{s. t. } h(W) = 0 \end{aligned} \quad (2)$$

式中,  $W$  是给定数据集,  $F(W)$  是与数据有关的损失函数,  $h(W)$  是一个平滑函数。

对于此非凸优化问题,可以将上式中的无环约束表示为

$$h(W) = \text{tr}(W \circ W) - d = 0 \quad (3)$$

式中,  $\text{tr}(\cdot)$  表示求矩阵的迹,“ $\circ$ ”表示 Hadamard 积。  $h(W)$  的梯度值为

$$\nabla h(W) = (e^{W \circ W})^T \circ 2W \quad (4)$$

得到最终需要进行优化的数学表达式:

$$\begin{aligned} \min_{W \in \mathbb{R}^{d \times d}} F(W) + \frac{\rho}{2} |h(W)|^2 \\ \text{s. t. } h(W) = 0 \end{aligned} \quad (5)$$

式中:  $\rho$  是大于 0 的一个二次惩罚项,以表示惩罚违反约束  $h(W) = 0$ 。至此可以采用 NO TEARS 算法对问题进行求解,具体步骤如算法 1 所示。

#### 算法 1 NO TEARS 算法

Alg. 1 NO TEARS algorithm

输入: 初始化解  $(w_0, \alpha_0)$ , 最小化速度  $c \in (0, 1)$ , 优化准确率  $\varepsilon > 0$ , 惩罚增长率  $\gamma > 1$

对  $t = 0, 1, 2, \dots$

(1) 对初始问题进行求解

$$W_{t+1} \leftarrow \underset{W}{\text{argmin}} L^p(W, \alpha_t)$$

(2) 若  $h(W_{t+1}) \geq c \cdot h(W_t)$ , 设定  $\rho \leftarrow \gamma\rho$ , 然后返回(1)

(3) 若  $h(W_{t+1}) < \varepsilon$ , 则停止

(4) 若  $h(W_{t+1}) \geq \varepsilon$ , 则使用对偶梯度上升法  $\alpha_{t+1} \leftarrow \alpha + \rho h(W_{t+1})$ , 重复上述步骤, 当满足  $h(W_{t+1}) < \varepsilon$  时停止

### 1.3 数据处理

通过分析 2018 年 1 月 20 日至 3 月 20 日的高速列车实绩运行数据和列车时刻表数据,得到高速列车延误信息,这些信息包括列车车次、日期、到达站点、图定到站时间、实际到站时间、延误时长(见表 1)。考虑到海南省和台湾省高铁站点暂时未与其他省份高速铁路网络相连接,所以本文只将上述两者以外省份的高速铁路网络作为研究范围(共 543 个站点, 481 523 条列车延误信息)。

表1 高速列车晚点信息示例  
Tab.1 Example train delay information

车次	日期	到达站点	图定到站时间	实际到站时间	延误时长/min
G1935	2018-01-20	上海虹桥	21:25	21:43	18
D2993	2018-02-05	桂林北	10:18	10:30	12
G71	2018-03-05	长沙南	14:26	15:22	56
G119	2018-03-20	南京南	14:20	14:47	27

1.4 模型构建

将高速列车主要运行时段 6:00—24:00 分割为 18 个 1 h 的时间窗,研究时段内共有 1 080 个时间窗。通过式(6)计算各站点在各时间窗内所有延误列车的平均延误时长,最后将其整理为 NO TEARS 贝叶斯网络结构学习算法输入所需的数据矩阵形式,即 1 080 行、543 列的站点平均延误时长矩阵  $T$ 。

$$T_{ij} = \sum_{n=1}^{N_{ij}} t_{ijn}^d / N_{ij}, T_{ij} \in T \quad (6)$$

式中,  $T_{ij}$  为  $i$  站点在  $j$  时间窗的列车平均延误时长,  $N_{ij}$  为  $i$  站点在  $j$  时间窗的延误列车数量,  $t_{ijn}^d$  为  $i$  站点在  $j$  时间窗内延误列车记录序号为  $n$  的列车延误时长。

根据 1.2 节推导过程,建立高速铁路网络站点延误传播贝叶斯网络模型的目标函数和约束条件:

$$\begin{aligned} \min_{W \in \mathbf{R}^{d \times d}} F(W; T) + \frac{\rho}{2} |h(W)|^2 \\ \text{s.t. } h(W) = 0 \end{aligned} \quad (7)$$

进一步使用 Python 3.8 集成环境 Anaconda,同时应用 NO TEARS 贝叶斯网络结构学习算法对应的 causalnex 贝叶斯网络库建立高速铁路网络站点延误传播贝叶斯网络,并分别采用 causalnex 贝叶斯网络库、pymPy 贝叶斯网络库和 Networkx 复杂网络库中有向无环图检验模块对所建立的网络结构进行检验,结果显示所建立的贝叶斯网络均通过算法的有向无环图检验。模型结构示意如图 1 所示。

2 基于渗流理论的延误传播规律分析

在复杂网络中,渗流理论<sup>[15]</sup>描述了网络中节点或边被移除时,网络结构发生变化的过程。当网络中某一参数发生变化,网络会出现巨大团簇分裂成若干个小团簇的现象;参数变化到某一临界阈值(渗流阈值)时,网络性质会突然发生改变,这种突变称为渗流突变。



图1 站点延误传播贝叶斯网络示意图  
Fig.1 Schematic diagram of delay propagation Bayesian network

2.1 延误传播团簇演变分析

站点延误传播贝叶斯网络模型中边和边的权重由 NO TEARS 算法获取,模型中的边表示两个站点之间的延误传播关系,边权重  $w_{ij}$  越大表示父节点站点  $i$  对子节点站点  $j$  造成的延误传播影响越强。基于渗流理论和已构建的站点延误传播贝叶斯网络,本文设置不同的权重阈值  $q$  观察站点延误传播贝叶斯网络的演变情况<sup>[16]</sup>。对于给定的权重阈值  $q$ ,站点延误传播贝叶斯网络中各条边根据权重  $w_{ij}$  被划分为有效边( $e_{ij} = 1$ )和无效边( $e_{ij} = 0$ ),如式(8)所示。网络中的站点通过有效边相互连接形成团簇,同一团簇中的站点存在有效边互连,不同团簇中的站点之间没有边相连。所以在给定权重阈值  $q$  下,根据不同团簇之间是否存在有效边连接,网络被划分为多个不同大小的延误传播团簇。

$$e_{ij} = \begin{cases} 1 & w_{ij} \geq q \\ 0 & w_{ij} < q \end{cases} \quad (8)$$

随着权重阈值  $q$  的变化,网络发生渗流突变过程。为了展示网络的渗流突变过程,图 2 展示了在不同权重阈值下网络中延误传播团簇的地理位置分布情况(分别选取小于、等于和大于渗流临界权重阈值的权重  $q = 0.2, q = 0.35, q = 0.6$ )。结合图 3 中不同权重阈值下的延误传播团簇规模

可以看出,随着权重阈值  $q$  逐渐增大,站点延误传播贝叶斯网络中最大延误传播团簇不断变小,而第二大延误传播团簇规模呈现震荡变大的趋势;当权重阈值  $q$  增加到渗流临界权重阈值  $q_c$  时,第二大延误传播团簇的规模达到峰值,而最大延误传播团簇突然瓦解,此时站点延误传播贝叶斯网

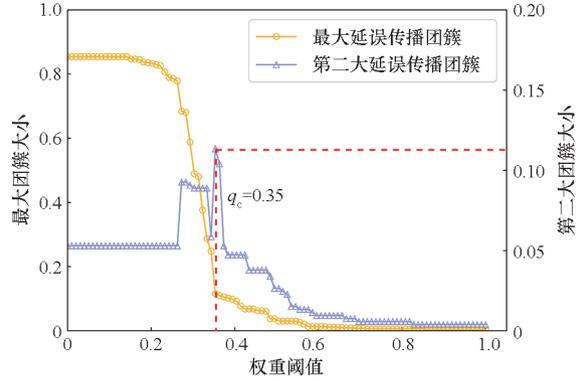


图 3 不同权重阈值  $q$  下延误传播团簇规模  
Fig. 3 Size of delay propagation clusters under different weight threshold  $q$  values

络处于渗流突变状态<sup>[15]</sup>;当权重阈值  $q$  大于渗流临界权重阈值  $q_c$  时,整体网络继续瓦解,直至网络全部解体。

### 2.2 渗流突变状态下的网络分析

在渗流理论中,网络突变十分关键。Wu 等<sup>[16]</sup>通过分析权重阈值等于渗流阈值的网络结构,发现核心团簇可以以较少的边展现原网络的骨架结构,体现出原网络的核心特征。在其他多数研究中<sup>[17-20]</sup>也将网络的权重阈值设置为渗流阈值,进而对网络中的核心团簇进行分析,从复杂的网络结构中找到重要的边。本节为寻找到网络中重要的延误传播关系,揭示网络中延误传播的核心特征,进一步分析图 2(b)中渗流突变状态下站点延误传播贝叶斯网络形成的核心延误传播团簇( $q_c = 0.35$ )。

度是网络图中描述节点属性的重要指标。节点  $i$  的度  $k_i$  定义为网络中与节点  $i$  直接相连的边的数量。在有向网络中,节点的度包括出度和入度。节点  $i$  的出度  $k_i^{out}$  定义为从节点  $i$  指向其他节点的边的数量,节点  $i$  的入度  $k_i^{in}$  定义为从其他节点指向节点  $i$  的边的数量<sup>[21]</sup>。

站点延误传播贝叶斯网络模型以有向边连接具有延误传播关系的两个站点,因此,在延误传播团簇中,根据站点的出度和入度,可将站点划分为 3 类:

1) 当  $k_i^{in} < k_i^{out}$  (站点的入度小于出度) 时,此类站点表现出主要将延误传递给其他站点的延误传播特性,这类站点归类为延误发散站点。

2) 当  $k_i^{in} = k_i^{out}$  (站点的入度等于出度) 时,此类站点表现出在受到延误传播影响后,站点具有较大可能将延误继续传播至其他站点的延误传播特性,这类站点归类为延误传递站点。

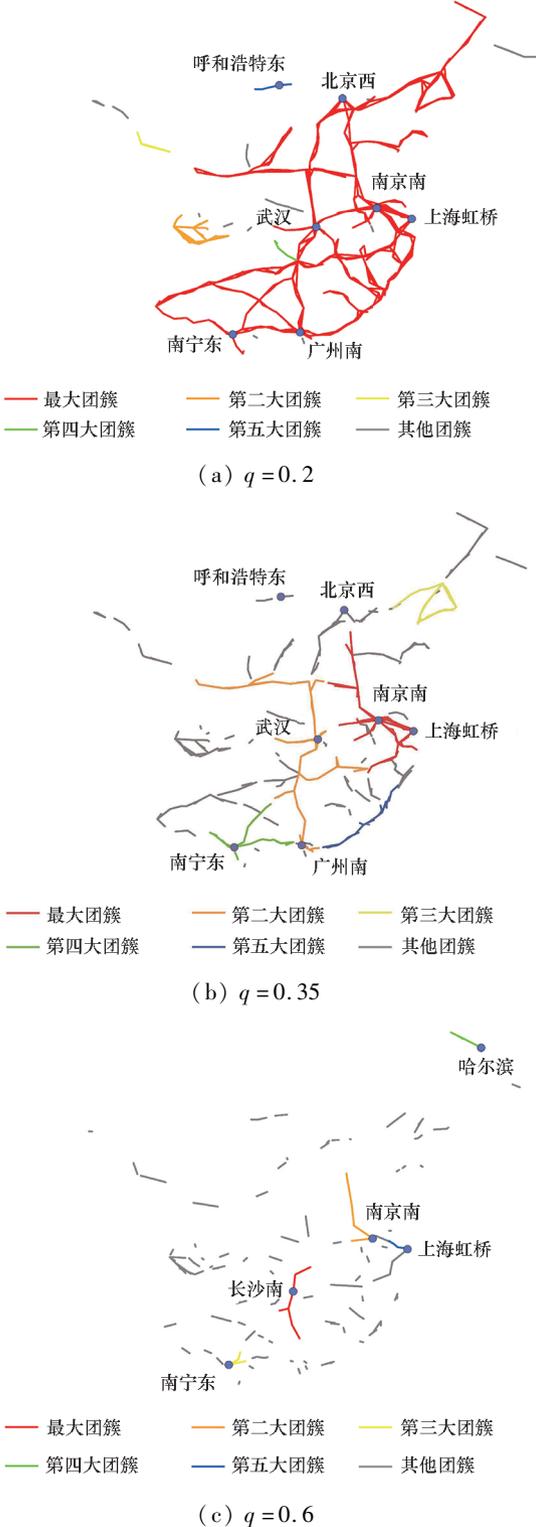


图 2 不同权重阈值  $q$  下延误传播团簇的地理位置  
Fig. 2 Delay propagation clusters' locations under different weight threshold  $q$  values

3) 当  $k_i^{in} > k_i^{out}$  (站点的入度大于出度) 时, 此类站点表现出主要受到其他站点延误传播影响的延误传播特性, 这类站点归类为延误消散站点。

南京南站在最大延误传播团簇内被划分为延误发散站点。图4展示了南京南站产生的延误传播链, 从图中可以看出南京南站可以首先将延误传播至附近的合肥南站、蚌埠南站等站点, 而合肥南站、蚌埠南站可以进一步将延误分别传播至六安站、徐州东站等站点; 在延误传播的过程中, 湖州站、徐州东站、济南西站等延误发散站点具有多条延误传播链, 进而可以在多方向将延误传播至浙江、河北等与南京南站相距较远地区的站点。这种延误传播模式使大部分延误发散站点成为延误传播中心(如图4中南京南站、济南西站、徐州东站等站点), 将延误“辐射式”传播至其他站点。

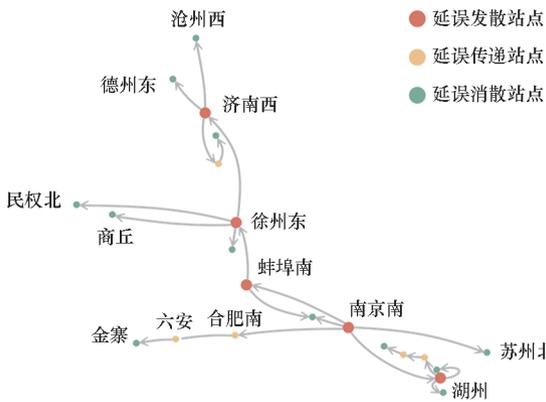


图4 南京南站产生的延误传播链

Fig. 4 Delay propagation chain generated by Nanjingnan Railway Station

其他延误传播团簇具有相似的延误传播规律。以第二大延误传播团簇为例, 图5展现了湖南地区的站点延误传播规律, 可以看出湖南地区

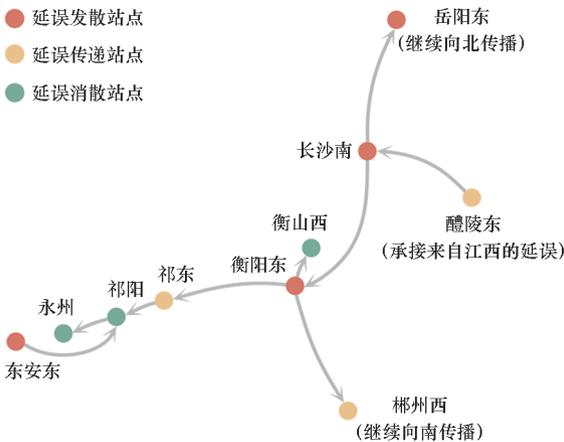


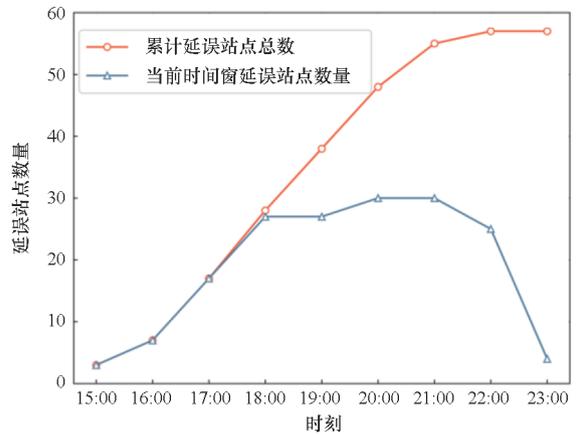
图5 长沙南站产生的延误传播链

Fig. 5 Delay propagation chain generated by Changshanan Railway Station

延误传播以长沙南站和衡阳东站这类延误发散站点为延误传播中心, 将延误以延误传播链的形式“辐射式”传播至其他站点。其中长沙南站向北可以将延误经岳阳东站继续传播至湖北地区, 向南可以传播至衡阳东站; 而衡阳东站也具有多条延误传播链, 可以将延误传播至祁东站、郴州西站等站点, 进而传播至永州和广东地区, 向网络继续扩散延误。

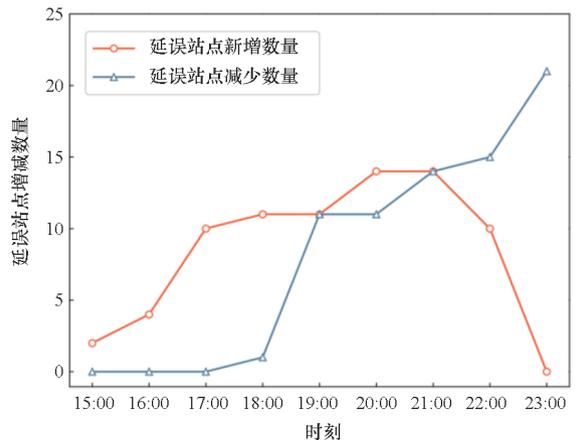
为从定量的角度分析网络中的延误传播规律, 本文进一步分析了受到延误发散站点延误传播影响的站点数量变化情况。在延误传播的过程中, 如果站点  $j$  受到站点  $i$  延误传播影响, 那么存在一条从站点  $i$  至站点  $j$  的传播路径满足以下要求: 延误由站点  $i$  开始, 随着时间窗的推移, 其路径上的站点沿传播路径方向依次发生延误。

图6展示了2018年3月15日网络中受到长沙南站延误传播影响的站点数量变化情况。从图



(a) 延误站点数量变化

(a) The number of delayed stations



(b) 延误站点增减数量变化

(b) Changes of the number of delayed stations

图6 受延误传播影响的站点数量  
Fig. 6 The number of stations affected by delay propagation

中可以看出:在 15:00—18:00 时间段,延误处于初始传播阶段,此时新增的延误站点数量不断增加,延误站点减少数量一直为 0,致使此时段内每一时间窗的延误站点数量和累计延误站点总数急剧升高。在 18:00—21:00 时间段,新增的延误站点数量和延误站点减少数量接近平衡,所以各时间窗的延误站点数量基本维持不变;但由于新增的站点数仍然保持在较高值,所以累计延误站点总数仍然急剧升高。在 21:00—23:00 时间段,由于大部分延误的高速列车在晚间结束全天的运行,延误在网络中扩散被阻碍,所以新增的延误站点数量逐步减少而延误站点总数继续增长,大部分站点从延误状态逐步恢复为正常状态。

### 3 站点延误状态预测

#### 3.1 预测模型建立

基于渗流突变状态下站点延误传播贝叶斯网络中的核心延误传播团簇,对高速铁路站点延误状态进行预测。为预测下一时间窗目标站点的延误状态,在站点延误状态预测模型中不仅考虑了目标站点的父节点站点延误状态,而且加入了两站点间的列车运行信息,以捕捉延误随列车运行传播到站点的动态变化情况。用于预测站点延误状态的贝叶斯网络模型结构如图 7 所示(在本图的案例中,待预测站点有两个父节点),其中模型变量集确定如下:

$$Y = \{S_{1j}, \dots, S_{nj}, S_{A_{j+1}}, Z_{1_{j+1}}, \dots, Z_{n_{j+1}}\}$$

式中,  $S_{A_{j+1}}$  为待预测站点  $A$  在  $j+1$  时间窗的延误状态,  $S_{nj}$  为站点  $A$  的第  $n$  个父节点站点在  $j$  时间窗的延误状态。中国铁路总公司将小于等于 4 min 的列车到站晚点视为可被车站运营组织管理所吸收的晚点,不会对列车和车站造成大的影响<sup>[22]</sup>。因此根据式(9)将站点的延误状态划分为延误状态  $S_{ij} = 1$  和非延误状态  $S_{ij} = 0$ 。

$$S_{ij} = \begin{cases} 1 & T_{ij} > 4 \\ 0 & T_{ij} \leq 4 \end{cases} \quad (9)$$

式中,  $S_{ij}$  为  $i$  站点在  $j$  时间窗的延误状态,  $T_{ij}$  为  $i$  站点在  $j$  时间窗的列车平均延误时长。

$Z_{n_{j+1}}$  表示在  $j+1$  时间窗内,站点  $A$  的第  $n$  个父节点站点与站点  $A$  之间是否有列车运行。因为两站点间并不是在所有时段内都会有高速列车运行,所以在预测站点延误状态时加入两站点间的列车运行信息  $Z_{n_{j+1}}$ ,以捕捉列车的运行动态变化情况。 $Z_{n_{j+1}}$  根据列车时刻表数据和  $j$  时间窗的列车运行信息得到,取值为  $\{1, 0\}$ ,其中 1 表示两

站点间有高速列车运行,0 表示无高速列车运行。

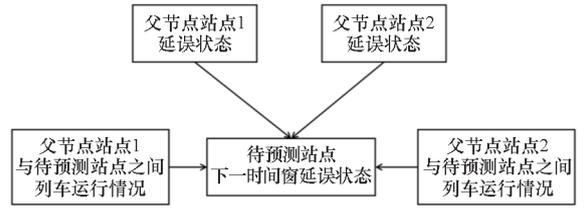


图 7 基于核心延误传播团簇的站点延误状态预测模型示意图

Fig. 7 Stations' delay state prediction model based on core delay propagation clusters

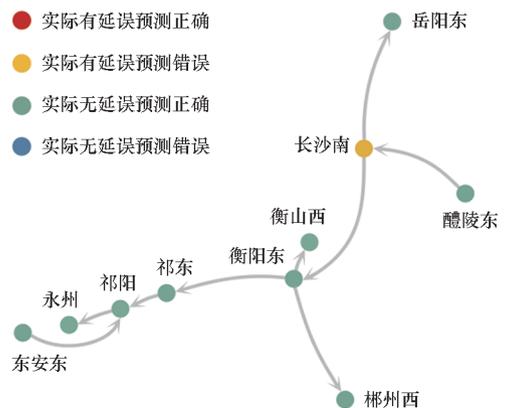
#### 3.2 预测结果分析

将 60 天数据集划分为模型的训练集和测试集,以数据集前 70% 时段即 2018 年 1 月 20 日至 2018 年 3 月 2 日的数据作为模型的训练集,以数据集后 30% 时段即 2018 年 3 月 3 日至 2018 年 3 月 20 日的数据作为模型的测试集。预测准确率  $R$  (计算见式(10)) 为 89.56%,表明基于核心延误传播团簇的预测模型很好地捕捉了网络中核心的延误传播关系。

$$R = \frac{y_{true}}{N} \quad (10)$$

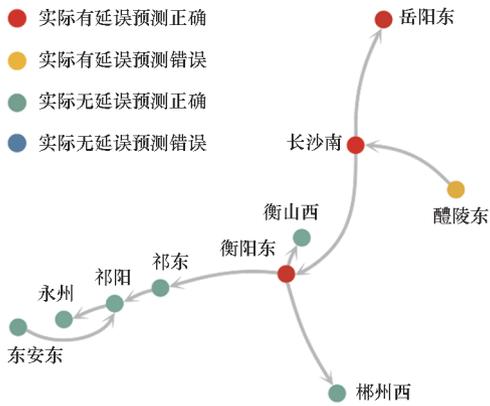
式中,  $N$  为预测样本数,  $y_{true}$  为正确预测站点是否处于延误状态的样本数。

为进一步说明模型的预测结果和网络中的延误传播规律,选取 2018 年 3 月 15 日的实际案例,对延误在网络中大范围传播的动态过程进行分析。图 8 展示了延误首先在长沙南站逐步传播至湖南地区的动态变化和模型的预测结果。从图中可以看出,14:00 湖南地区长沙南站首先处于延误状态,附近的其他站点此时都处于非延误状态;随着时间的推移,15:00 长沙南站的延误已经传播到了岳阳东站、衡阳东站,进而形成



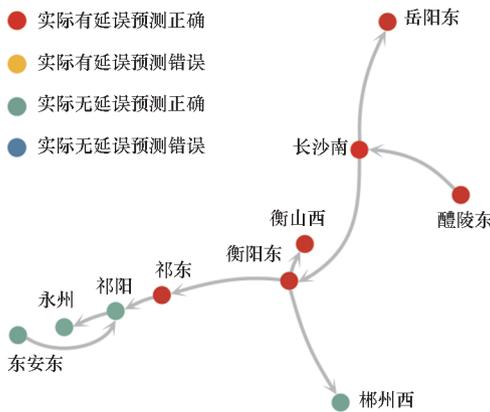
(a) 14:00 预测结果

(a) Prediction result at 14:00



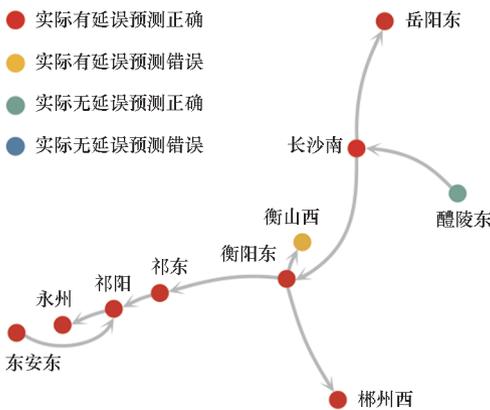
(b) 15:00 预测结果

(b) Prediction result at 15:00



(c) 16:00 预测结果

(c) Prediction result at 16:00



(d) 17:00 预测结果

(d) Prediction result at 17:00

图 8 湖南地区预测结果示例

Fig. 8 Example prediction result in Hunan province

多条延误传播链,在后续时段内将延误传播至郴州西站、衡山西站等更多站点;直到 17:00,湖南地区大部分站点都处于延误状态。预测模型虽然未能捕捉到 14:00 长沙南站产生的突发性延误,但能够精准地预测延误的动态传播趋势,对于网

络中站点延误预警有着重要意义。

### 4 结论

本文建立了高速铁路网络的站点延误传播贝叶斯网络模型,并结合渗流理论对高速铁路网络延误传播规律进行研究,主要得到以下结论:随着权重阈值  $q$  增大,站点延误传播贝叶斯网络中最大延误传播团簇不断减小,而第二大延误传播团簇在渗流临界权重阈值  $q_c$  时达到最大;从延误传播的角度将站点划分成延误发散站点、延误传递站点和延误消散站点,其中延误发散站点产生的延误传播链使网络以部分延误发散站点为延误中心向延误传递站点和延误消散站点进行“辐射式”延误传播。在预测站点延误状态方面,由于延误传播团簇很好地捕捉了网络中核心的延误传播关系,能够用于站点延误状态预警。

研究结果可以帮助管理者从全局性视角理解延误在网络中的传播机理,进而制定更高效的高速铁路行车组织管理方案,以防止延误从局部网络向全网扩散。未来可在现有基础上融合多源数据更深层次地分析站点发生延误的原因,并结合专家领域知识提升模型预测精度。

### 参考文献 (References)

- [1] FLIER H, GELASHVILI R, GRAFFAGNINO T, et al. Mining railway delay dependencies in large-scale real-world delay data [ M ]// Robust and Online Large-Scale Optimization. Berlin, Heidelberg: Springer, 2009: 354 - 368.
- [2] 方劲劼,朱昌锋,成琳娜,等. 基于渗流理论的铁路列车延误传播模型研究[J]. 兰州交通大学学报, 2023, 42(1): 28 - 37.  
FANG J H, ZHU C F, CHENG L N, et al. Railway delay propagation model based on method of percolation theory[J]. Journal of Lanzhou Jiaotong University, 2023, 42(1): 28 - 37. (in Chinese)
- [3] MAGADAGELA K, NEL H, MARNEWICK A. Identification of delay factors that affect high dwell times of freight trains[C]//Proceedings of IEEE Technology & Engineering Management Conference (TEMSCON), 2017: 179 - 184.
- [4] WEN C, HUANG P, LI Z C, et al. Train dispatching management with data-driven approaches: a comprehensive review and appraisal[J]. IEEE Access, 2019, 7: 114547 - 114571.
- [5] MEESTER L E, MUNS S. Stochastic delay propagation in railway networks and phase-type distributions [ J ]. Transportation Research Part B: Methodological, 2007, 41(2): 218 - 230.
- [6] 胡思继,孙全欣,胡锦涛,等. 区段内列车晚点传播理论的研究[J]. 中国铁道科学, 1994, 15(2): 41 - 54.  
HU S J, SUN Q X, HU J Y, et al. Research on theories of train delay propagation in a railway district [ J ]. China Railway Science, 1994, 15(2): 41 - 54. (in Chinese)

- [7] 庄河, 文超, 李忠灿, 等. 基于高速列车运行实绩的致因-初始晚点时长分布模型[J]. 铁道学报, 2017, 39(9): 25-31.  
ZHUANG H, WEN C, LI Z C, et al. Cause based primary delay distribution models of high-speed trains on account of operation records[J]. Journal of the China Railway Society, 2017, 39(9): 25-31. (in Chinese)
- [8] WANG P, ZHANG Q P. Train delay analysis and prediction based on big data fusion [J]. Transportation Safety and Environment, 2019, 1(1): 79-88.
- [9] 柏赞, 何天健, 毛保华. 一种交叉线干扰情形下列车晚点恢复运行控制方法[J]. 交通运输系统工程与信息, 2011, 11(5): 114-122.  
BAI Y, HE T J, MAO B H. Train control to reduce delays upon service disturbances at railway junctions[J]. Journal of Transportation Systems Engineering and Information Technology, 2011, 11(5): 114-122. (in Chinese)
- [10] 曾壹, 张琦, 陈峰. 一种实时抗干扰的列车节能-晚点恢复模型[J]. 交通运输系统工程与信息, 2018, 18(2): 171-176.  
ZENG Y, ZHANG Q, CHEN F. An energy-delay optimization method for real-time train path disturbances in transportation network[J]. Journal of Transportation Systems Engineering and Information Technology, 2018, 18(2): 171-176. (in Chinese)
- [11] LESSAN J, FU L P, WEN C. A hybrid Bayesian network model for predicting delays in train operations[J]. Computers & Industrial Engineering, 2019, 127: 1214-1222.
- [12] CORMAN F, KECMAN P. Stochastic prediction of train delays in real-time using Bayesian networks [J]. Transportation Research Part C: Emerging Technologies, 2018, 95: 599-615.
- [13] 许保光, 王蓓蓓, 池宏, 等. 基于贝叶斯网络的航空安全中不安全信息分析[J]. 中国管理科学, 2020, 28(12): 118-129.  
XU B G, WANG B B, CHI H, et al. Insecurity information analysis in civil aviation safety based on Bayesian network[J]. Chinese Journal of Management Science, 2020, 28(12): 118-129. (in Chinese)
- [14] ZHENG X, ARAGAM B, RAVIKUMAR P, et al. DAGs with NO TEARS: continuous optimization for structure learning[C]// Proceedings of the 32nd International Conference on Neural Information Processing Systems, 2018: 9492-9503.
- [15] SHANTE V K S, KIRKPATRICK S. An introduction to percolation theory[J]. Advances in Physics, 1971, 20(85): 325-357.
- [16] WU Z H, BRAUNSTEIN L A, HAVLIN S, et al. Transport in weighted networks: partition into superhighways and roads[J]. Physical Review Letters, 2006, 96(14): 148702.
- [17] LI D Q, FU B W, WANG Y P, et al. Percolation transition in dynamical traffic network with evolving critical bottlenecks[J]. Proceedings of the National Academy of Sciences, 2015, 112(3): 669-672.
- [18] LI S N, YANG H, LI M L, et al. A highway on-ramp control approach integrating percolation bottleneck analysis and vehicle source identification [J]. Sustainability, 2023, 15(16): 12608.
- [19] LYU S R, YANG H, ZHANG F, et al. Identifying the bottlenecks of urban metros and analyzing the passenger source[C]// Proceedings of 22nd COTA International Conference of Transportation Professionals, 2022: 2625-2635.
- [20] MOLINERO C, MURCIO R, ARCAUTE E. The angular nature of road networks [J]. Scientific Reports, 2017, 7: 4312.
- [21] 汪小帆, 李翔, 陈关荣. 网络科学导论[M]. 北京: 高等教育出版社, 2012.  
WANG X F, LI X, CHEN G R. Network science: an introduction[M]. Beijing: Higher Education Press, 2012. (in Chinese)
- [22] 黄平, 彭其渊, 文超, 等. 高速铁路故障分类及其影响列车数模型[J]. 中国安全科学学报, 2018, 28(增刊2): 46-53.  
HUANG P, PENG Q Y, WEN C, et al. Study on high-speed railway disruption classification and model of its influence on train number [J]. China Safety Science Journal, 2018, 28(Suppl 2): 46-53. (in Chinese)