

文章编号: 2095-2163(2024)02-0124-04

中图分类号: U491

文献标志码: A

# 基于机器学习的交通网络级联失效态辨识与预测

张腾飞, 郑丽, 郝龙, 李玉龙

(兰考三农职业学院 旅游服务与管理系, 河南 开封 475300)

**摘要:** 随着交通路网拥堵的问题加剧, 道路交通拥堵又会演变成交通网络级联失效的问题。为有效缓解交通堵塞造成的级联失效问题, 准确的短时交通流预测与级联失效态预测可以为交通路网提供可靠预警信息。首先, 针对机器学习训练样本不精确的情况, 提出了基于 KNN-BPNN 交通流预测模型, 对短时交通流数据进行在线预测; 其次, 用机器学习的 K 均值聚类方法和 KNN 算法对交通状态进行判别, 在此基础上给出级联失效态的识别方法对级联失效态进行识别; 最后, 提出基于 BPNN 的交通网络级联失效态模型进行预测。实例分析验证模型的结果表明, 本文提出的 KNN-BPNN 组合预测模型可以为级联失效态预测提供可靠的交通流数据信息; BPNN 在交通网络级联失效态预测方面也表现出较好的预测精度, 可以准确的为道路交通网络级联失效提供预警信息。

**关键词:** 机器学习; 神经网络; K 均值聚类; KNN; 交通网络; 级联失效态

## Identification and prediction on cascading failure state in transportation networks based on machine learning

ZHANG Tengfei, ZHENG Li, HAO Long, LI Yulong

(Department of Tourism and Management, Lankao Vocational College of San Nong, Kaifeng Henan 475300, China)

**Abstract:** As traffic congestion in road networks worsens, it escalates into the issue of cascading failures in the transportation network. To effectively alleviate this problem, accurate short-term traffic flow prediction and cascading failure state prediction can provide reliable warning information for the transportation network. To address the challenge of imprecise machine learning training samples, we propose a KNN-BPNN traffic flow prediction model for online short-term traffic flow predictions. We also employ machine learning techniques, such as K-means clustering and KNN algorithms, to discern traffic states, enabling the identification of cascading failure states. Furthermore, we introduce a BPNN-based model for predicting cascading failure states within the transportation network. Experiments validate the results, demonstrating that the KNN-BPNN combined prediction model provides dependable traffic flow data for predicting cascading failure states. BPNN also exhibits strong predictive accuracy in forecasting transportation network cascading failure states, offering accurate warning information for road traffic network cascading failures.

**Key words:** machine learning; neural network; K-means clustering; KNN; traffic network; cascading failure state

## 0 引言

城市交通网络中引起拥堵的原因主要有交通事件、自然恶劣环境、人为蓄意破坏或高峰期交通堵塞等, 可能会进一步引起交通网络的级联失效。对城市交通网络级联失效进行研究, 不仅要考虑到路网基础设施, 更应该从交通出行网络的交通流角度对级联失效态做出准确的辨识与预测。在当前大数据背景下, 海量的交通数据也越来越多样化, 通过机器学习的方法对交通数据进行处理成为研究的主流。

目前国内外关于城市交通网络级联失效态的相关研究主要集中在级联失效过程建模和级联失效的影响, 王杰<sup>[1]</sup>考虑交通拥挤的传递性, 建立双层规划模型, 探讨不同节点删除方法下关键节点的识别, 提出 3 种能够有效控制城市路网级联失效的策略; Qian 等<sup>[2]</sup>利用对偶方法, 在道路交通网络中加入时间延迟和消散系数来模拟小世界特征的道路交通网络级联失效, 采用事件消散系数和容量系数合理地避免级联失效; 李成兵等<sup>[3]</sup>研究多种运输方式的加权网络模型, 分析得出连续两次蓄意攻击会使级联失效更

**基金项目:** 河南省高等学校重点科研项目(23B460019)。

**作者简介:** 郑丽(1991-), 女, 硕士, 助教, 主要研究方向: 职业教育; 郝龙(1989-), 男, 硕士, 讲师, 主要研究方向: 大数据分析; 李玉龙(1989-), 男, 硕士, 讲师, 主要研究方向: 智能制造装备设计。

**通讯作者:** 张腾飞(1994-), 男, 硕士, 助教, 主要研究方向: 交通运输。Email: 1165413352@qq.com

**收稿日期:** 2023-02-17

严重;邢茹茹<sup>[4]</sup>建立了偶发性拥挤的级联失效模型和常发性拥挤的级联失效模型,并提出了限流的级联失效控制模型,给出控制级联失效优化策略。

综上所述,目前国内外学者关于交通网络级联失效的研究已有些许成果,在交通网络级联失效态的识别与预测方面还鲜有相关研究。本文通过机器学习的方法对级联失效态的辨识与预测展开研究,首先,从交通流预测的角度建立基于 KNN-BPNN 模型对交通数据预测;其次,以无监督学习 K 均值聚类方法和有监督 KNN 算法对交通状态进行分类判别,得到级联失效态数据;最后,通过 BPNN 模型对交通网络级联失效态预测。

### 1 模型基础理论

机器学习是指利用计算机等工具模拟大自然学习的一种学习过程,主要是从大量的训练数据中挖掘出分类识别或者回归预测的决策函数,对未知数据进行判别。机器学习不断的发展,研究方向也不断被细化,按机器学习的样本学习方法分类,可以分为三类:无监督学习、半监督学习和有监督学习<sup>[5]</sup>。经典的机器学习有神经网络、K 近邻、支持向量机、K 均值聚类等算法。

误差反向传播神经网络(BPNN)。BPNN 在学习大数据和存储规律上有很大的优势,可以学习到大数据内部的映射规律且存储在自身记忆中,无须通过数学知识表示出隐藏的规律,通过历史数据训练可以得到数据潜在规律。

K 近邻算法(K-Nearest Neighbor, KNN)不仅有着较好的分类预测能力,还可以用于回归预测。KNN 算法是经典的机器学习方法,计算时需要考虑预测样本数据与所有历史数据之间的距离,选择最相近的数据作为预测数据的决策样本,在回归预测中有较好的预测精度<sup>[6]</sup>。

K 均值聚类算法是通过距离计算样本之间相似度的分类算法,是机器学习中经典的无监督学习算法,因其快速、高效、简便的特性而被广泛应用于聚类分析中。

#### 1.1 KNN-BPNN 模型

BPNN 是比较经典的机器学习模型,在短时交通流预测中可以识别出短时交通流中存在的非线性特征,更好的拟合数据,得到较优的预测模型。K 近邻算法作为机器学习的经典算法,在数据挖掘与搜索方面有特定的优势,在回归预测中可以发挥数据搜索能力,得到所有与待测样本比较相近的历史数据。

神经网络模型在数据训练时会把全部历史数据作为输入,通过神经网络的有监督训练得到非参数模型,进一步对交通流数据预测。然而神经网络虽能很好的拟合时间序列的非线性特征,但是短时交通流数据中会有一些非线性特征会对预测产生负面的影响,所以选取优质的训练数据对模型的预测性能起着关键的作用,有必要对神经网络的训练数据进行严格的把控。而 K 近邻算法是一种基于样本距离计算的机器学习方法,在进行数据搜索方面有自己得天独厚的优势,可以在全部输入数据中搜索得到合适的部分数据作为神经网络的输入数据,采取高效的数据训练神经网络,势必会带来更好的预测性能。

本文创新性的提出一种基于 KNN-BPNN 的组合预测模型,充分利用 K 近邻算法的数据搜索能力和 BPNN 模型的非线性拟合能力,以 K 近邻搜索最合适的训练样本作为 BPNN 的数据,建立预测模型,对短时交通流数据进行在线预测。首先,通过 KNN 算法对 BP 神经网络的训练数据进行时间序列的重构与简化;将得到的样本数据作为 BPNN 的输入,训练模型;最后,对实时交通流数据预测。通过此组合模型可以提高 BPNN 模型的预测性能,KNN-BPNN 的建模流程如图 1 所示。

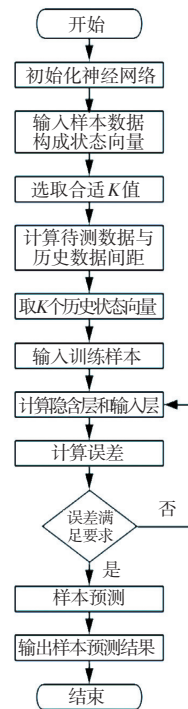


图 1 KNN-BPNN 建模流程

Fig. 1 Pipeline of KNN-BPNN

#### 1.2 交通网络级联失效态

城市交通网络是一种复杂的双层网络,是由道

路网络结构层和交通出行网络层构成的复杂网络。道路网络结构与出行网络之间是相互联系、相互影响的。在交通网络中,交通流量是不断变化的,当网络中的交叉口或路段发生失效或再现等变动时,道路网络的分布也会因为交通拥堵的变化而产生变化。

结合以往级联失效态的定义和实际道路交通网络,本文给出交通网络级联失效态定义:在交通路网中,由于突发拥堵导致部分关键路口或路段拥堵失效,交通出行者会根据路网拥堵情况和自身出行目的再次选择邻近的出行路线,交通流量又会再次传播到有效路段,由于流量的过度传播又会导致该路段拥堵失效,如此反复导致区域性路网失效的现象。

为了准确的判断不同路段在交通网络中是否失效,参考中国交通流相关理论,并结合本实验实际道路交通状况,本文采用交通流三参数即交通量、速度、密度,将上层交通网络的交通状态分为5类即自由流、顺畅流、平稳流、拥堵流和堵塞流<sup>[7]</sup>;下层道路网络在是否失效方面仅分为两种状态,包括失效态(拥堵流和堵塞流)和非失效态(自由流、顺畅流和平稳流)。

在交通网络中,判断是否发生级联失效的前提需要确定交通网络是否失效,而判断交通网络是否失效的前提是需要确定网络中各个路段是否失效。为此本文从路段是否发生失效为基础展开讨论,假设一个交通网络的下层道路网络有 $N$ 条路段,上层出行网络有 $M$ 种交通状态,记 $t$ 时刻路段 $i$ 是否失效的状态为 $L_i(t)$ ,式(1):

$$L_i(t) = \begin{cases} 1, K \in M \\ 0, J \in M, J \neq K \end{cases} \quad (1)$$

$L_i(t) = 1$ 表示路段 $i$ 为非失效态; $L_i(t) = 0$ 表示路段 $i$ 为失效态。整个路网的道路状态是否失效可以表示为 $L(t) = \{L_1(t), L_2(t) \dots L_N(t)\}$ 。

在一个路段数为 $N$ 的交通网络中,出行网络有 $M$ 种交通状态,按照失效态或非失效态两种状态可以算出该路网共有 $2^N$ 种状态。当 $t$ 时刻道路网络中有超过半数路段失效不能满足正常的通行需求,判定该时刻的道路网络处于失效态;否则 $t$ 时刻的路段能满足正常的通行需求,定义该路段处于非失效态。记 $t$ 时刻交通网络是否失效的状态为 $W(t)$ ,式(2):

$$W(t) = \begin{cases} 1, Q \in L \\ 0, P \in L, P \neq Q \end{cases} \quad (2)$$

其中, $P$ 和 $Q$ 表示整个路网中所有路段的状态变量; $W(t) = 1$ 表示该交通网络为非失效态; $W(t) = 0$ 表示该交通网络为失效态。

路网状态属于交通网络失效态或非失效态,并不是所有的网络失效都是属于级联失效态。有的网络失效是因为部分路段同时失效导致整个网络失效,这种网络失效在时间上没有因果关系,称为“共因失效态” $S(t) = 1$ ;有的网络失效,但是这些前后失效路段之间并没有直接的因果关系,这种网络失效称为“一般失效态” $S(t) = 1$ ;有的网络失效是某一时间交通网络中某一条或几条路段失效,进而导致下一时间与该失效路段相关的路段也随之失效,最终导致整个网络处于失效状态,这种失效不仅在时间上存在前后因果关系,在失效路段上也存在因果关系,这种失效状态称为“级联失效态” $S(t) = 0$ 。

## 2 实证研究

### 2.1 KNN-BPNN 的交通流预测

本文采用常州市16条路段的真实数据(取单向交通通行数据)组成路网做案例分析,数据时间为2021年10月7日至9日,每日0:00-2:00和5:30-24:00的短时交通流参数数据,采取数据包括交通量(辆/3分钟)、速度(千米/时)、密度(辆/千米),数据获取时间间隔为3分钟。

通过数据的预处理,得到以交通量为基础的各个状态向量的训练样本,预测时将 $t$ 时刻之前的5个连续时刻的数据作为KNN模型的输入,经过KNN的数据搜索得到与输入数据最相似的状态向量数据作为神经网络的训练样本,经过神经网络训练得到KNN-BPNN训练模型,对测验数据进行预测,以10月7日数据为训练样本对10月8-9日的的数据预测得到交通流预测数据。

### 2.2 交通网络级联失效态辨识

采用常州市2021年10月7日的交通流参数。首先,以K均值聚类方法对交通状态分类,将交通状态分为5类;其次,以KNN算法对2021年10月8-9日对交通状态判别,进一步可以进行级联失效态的辨识与预测。

交通网络级联失效态的辨识是对交通网络中各个路段状态组合模式的识别,本实验的交通网络级联失效态的辨识建立在一个简单的道路网络之上,数据采用常州市16条路段组成的简单路网,如图2所示。采用本文提出的级联失效态识别方法对2021年10月7-9日的真实交通状态数据进行交通

网络级联失效态的识别。

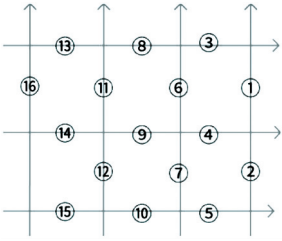


图 2 交通网络示意图

Fig. 2 Schematic diagram of traffic network

根据定义,在  $t$  时刻若有超过半数的路段失效时,认为网络失效,即当 16 条路段中有 9 条及以上路段失效时,认为路网失效,共有 26 333 种组合。 $t$  时刻的路网失效状态  $W(t) = 0$ , 进一步观察失效路段之间是否存在上下游之间的连接关系,若存在,就可以判断  $(t - 1)$  时刻中是否有失效路段;若  $(t - 1)$  时刻中存在失效路段,且与  $t$  时刻失效路段之间存在近邻上下游的连接关系,则判定  $t$  时刻的交通网络状态  $W(t)$  为级联失效态  $S(t) = 1$ , 即当  $t$  时刻和  $(t - 1)$  时刻的交通状态组合均满足上述条件时,判断  $t$  时刻属于级联失效态。得到级联失效态数据,就可将这些数据作为训练样本,建立级联失效态预测模型。

### 2.3 基于 BPNN 的交通网络级联失效态预测

本文采用 K 均值聚类得到的 10 月 7 日所有路段的交通状态和级联失效态数据作为训练样本,预测样本为 KNN 算法得到 10 月 8-9 日预测交通状态数据。用 BP 神经网络模型预测交通网络级联失效态,实验确定该 BP 神经网络有 33 个输入层节点,9 个隐含层节点以及 1 个输出层节点。对训练样本进行学习,得到 BPNN 预测模型,以 KNN 预测得到的交通状态数据为测试样本,以  $t - 1$  时刻的交通状态数据、级联失效态数据和  $t$  时刻判别得到的交通状态数据作为输入数据,将所有数据输入 BPNN 预测模型,得到  $t$  时刻预测的级联失效态,预测结果如图 3 和表 1 所示。

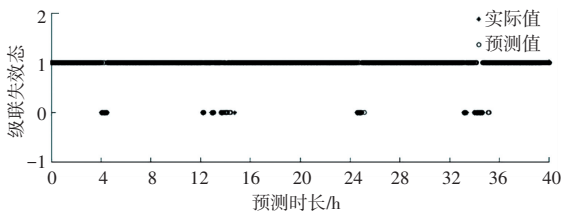


图 3 级联失效态的预测

Fig. 3 Prediction of cascading failure state

表 1 交通网络级联失效态的预测结果

Table 1 Prediction results of cascading failure state of traffic network

	实际级联失效态	预测级联失效态	准确率/%
发生次数	8	7	87.5
累计时间/min	162	135	83.3

从图 3 和表 1 的预测结果可知,实际交通状态数据的级联失效态共发生了 8 次,发生时间共计为 162 分钟。而预测得到的级联失效态包括实际级联失效态 7 次,发生时间 135 分钟。从级联失效态的发生时间计算,本文的预测准确率为 83.3%;从整个级联失效态的预测结果分析,预测准确率为 87.5%。说明通过机器学习的方法可以很好的预测短时交通网络级联失效态的发生,能够为智能交通系统提供可靠道路状态信息。

### 3 结束语

本文针对短时交通流预测问题提出了基于 KNN-BPNN 组合预测模型,给出了针对交通网络级联失效态的识别预测问题的相关定义,提出了基于 BPNN 的级联失效态预测模型。通过实例分析验证模型的结果表明,KNN-BPNN 组合预测模型可以充分发挥 K 近邻算法和 BP 神经网络模型的优点,为级联失效态预测提供可靠交通数据信息;BPNN 模型在交通网络级联失效态预测方面也表现出较好的预测精度,可以准确地为道路交通提供预警信息。本文研究方向侧重于级联失效态的预测,如何更便捷地进行交通网络级联失效态的识别将是今后的研究方向。

### 参考文献

- [1] 王杰. 城市道路交通网络级联失效及控制研究[D]. 长沙:长沙理工大学, 2014.
- [2] QIAN Y, WANG B, XUE Y, et al. A simulation of the cascading failure of a complex network model by considering the characteristics of road traffic conditions[J]. Nonlinear Dynamics, 2015, 80(1-2): 413-420.
- [3] 李成兵,张帅,杨志成,等. 蓄意攻击下城市群客运交通网络级联抗毁性仿真[J]. 交通运输系统工程与信息, 2019, 19(2): 14-21.
- [4] 邢茹茹. 城市区域路网级联失效主动防控理论及方法研究[D]. 长春:吉林大学, 2021.
- [5] 苗超维,秦品乐. 基于多分类 SVM 和 Hd 的目标跟踪算法[J]. 计算机工程与设计, 2016, 37(11): 3118-3123.
- [6] 郑智泉,王孟孟,田维琦. 基于加权 K 近邻算法的缺失数据填补研究[J]. 智能计算机与应用, 2021, 11(11): 31-33, 42.
- [7] 李爱青,于成功. 西宁市道路交通流运行拥堵特性研究[J]. 青海交通科技, 2018(3): 65-69.