

TaGNN: 基于趋势感知图神经网络的耦合噪声水质预测^{*}

孙建明^{1,2}, 徐宇扬¹, 仝 硕³, 应豪超³, 张 啸⁴, 庄福振⁵, 吴 健³

¹(浙江大学 计算机科学与技术学院, 浙江 杭州 310058)

²(苏州市伏泰信息科技股份有限公司, 江苏 苏州 215124)

³(浙江大学 公共卫生学院, 浙江 杭州 310058)

⁴(山东大学 计算机科学与技术学院, 山东 青岛 266237)

⁵(北京航空航天大学 人工智能学院, 北京 100191)

通信作者: 应豪超, E-mail: haochaoying@zju.edu.cn



摘 要: 未来水质预测任务根据各个观测节点的历史水质数据及相应的拓扑位置关系, 预测未来水质变化情况, 是图神经网络辅助环境保护的重要任务之一. 然而, 采集数据的数值和节点间的拓扑结构均存在噪声, 且噪声存在耦合现象. 同时, 污染物流向的不同导致数值噪声与结构噪声的耦合现象更为复杂难以解耦. 因此, 提出了一种基于趋势感知图神经网络的耦合噪声水质预测方法: 1) 利用历史水质数据趋势特征, 挖掘原始水质数据指标的局部相互关系, 构建多种潜在的水系拓扑结构, 分离结构噪声; 2) 利用构建的邻接矩阵与原始数据挖掘时空特征, 分离数值噪声. 最终, 根据潜在结构构建前后图网络节点表征一致性聚合水质预测结果. 实验结果表明, 所提方法在真实数据集上比当前最优方法表现更好, 且给出的水系潜在拓扑结构符合真实情况. 代码和数据可在 GitHub 上访问: <https://github.com/aTongs1/TaGNN>.

关键词: 水质预测; 图神经网络; 时空数据; 结构噪声; 图噪声学习

中图法分类号: TP18

中文引用格式: 孙建明, 徐宇扬, 仝硕, 应豪超, 张啸, 庄福振, 吴健. TaGNN: 基于趋势感知图神经网络的耦合噪声水质预测. 软件学报, 2025, 36(11): 5227–5240. <http://www.jos.org.cn/1000-9825/7381.htm>

英文引用格式: Sun JM, Xu YY, Tong S, Ying HC, Zhang X, Zhuang FZ, Wu J. TaGNN: Tendency-aware Graph Neural Network for Water Quality Prediction with Coupled Noise. Ruan Jian Xue Bao/Journal of Software, 2025, 36(11): 5227–5240 (in Chinese). <http://www.jos.org.cn/1000-9825/7381.htm>

TaGNN: Tendency-aware Graph Neural Network for Water Quality Prediction with Coupled Noise

SUN Jian-Ming^{1,2}, XU Yu-Yang¹, TONG Shuo³, YING Hao-Chao³, ZHANG Xiao⁴, ZHUANG Fu-Zhen⁵, WU Jian³

¹(College of Computer Science and Technology, Zhejiang University, Hangzhou 310058, China)

²(Suzhou Vortexinfo Technology Co. Ltd., Suzhou 215124, China)

³(School of Public Health, Zhejiang University, Hangzhou 310058, China)

⁴(School of Computer Science and Technology, Shandong University, Qingdao 266237, China)

⁵(School of Artificial Intelligence, Beihang University, Beijing 100191, China)

Abstract: The prediction of future water quality, which involves leveraging historical water quality data from various observation nodes and their corresponding topological relationships, is recognized as a critical application of graph neural networks in environmental protection. This task is complicated by the presence of noise within both the collected numerical data and the inter-node topological structures, compounded by a coupling phenomenon. The varying directions of pollutant flow intensify the complexity of coupling between

* 基金项目: 国家自然科学基金 (62106218); 浙江省自然科学基金 (LY23F020019)

收稿时间: 2024-05-23; 修改时间: 2024-10-24; 采用时间: 2024-12-17; jos 在线出版时间: 2025-04-23

CNKI 网络首发时间: 2025-04-23

numerical and structural noise. To address these challenges, a novel tendency-aware graph neural network is proposed for water quality prediction with coupled noise. First, historical water quality trend features are used to uncover local interdependencies among raw water quality indicators, enabling the construction of multiple potential hydrological topological structures and the disentanglement of structural noise. Second, spatio-temporal features are extracted from the constructed adjacency matrices and original data to separate numerical noise. Finally, water quality predictions are obtained by aggregating coherent node representations derived from the inferred latent structures across pre- and post-structure construction phases. Experimental results demonstrate that the proposed method outperforms state-of-the-art models on real-world datasets and generates potential hydrological topological structures that closely align with actual observations. The code and data are publicly available on GitHub: <https://github.com/aTongs1/TaGNN>.

Key words: water quality prediction (WQP); graph neural network (GNN); spatio-temporal data; structural noise; graph noise learning

城市水资源的监测保护是社会绿色可持续发展的重要举措. 浮标传感器自动采集、大坝断面处定期采集等多种方式实现了城市水质动态监测数字化. 海量多指标水质数据的挖掘分析, 对研判水资源变化趋势, 避免水质恶化引发次生环境伤害具有重大意义. 未来水质预测 (water quality prediction, WQP) 利用某一时间段的历史水质数据, 探寻未来一个或几个时间点的水质变化情况^[1], 是一类时间-空间预测任务 (spatio-temporal prediction). 该任务在水资源日常监测中具有重要的实际应用价值. 例如, 监测人员可以提前探测到河道污染风险, 向河道上下游站点发送预警信号, 追溯可能的污染源; 针对预测与实际情况差异较大的地点, 可快速启动人工巡查, 提升现场执法的效率, 避免大规模污染风险.

现有的水质预测研究往往利用图神经网络 (graph neural network, GNN) 与递归神经网络 (recurrent neural network, RNN) 对时空关系进行挖掘^[1-9]. 然而, 数据采集过程中存在的大量噪声问题尚未被充分关注. 首先, 观测数据存在数值误差. 浮标传感器频繁维护困难、温漂现象易发等客观因素导致观测数值存在固有误差和随机误差; 不同传感器采集数据频率不一导致观测数值在时间点上难以有效对齐. 其次, 观测点间的空间关系存在结构误差. 城市地下暗渠与隐蔽人工排口错综复杂, 各个观测点间的连通关系难以精确获取, 导致采集的空间数据存在大量结构噪声.

近几年考虑噪声问题的图神经网络鲁棒训练方法成为研究热点. 这些工作^[10-12]往往基于自监督或无监督学习对节点生成伪标签, 利用图数据中相同标签节点相连的特性和节点的邻居信息依赖性, 在原始图结构中引入额外边, 经过迭代对节点的数值噪声进行去噪. 此外, 也有一些工作^[13]利用结构熵等理论对噪声结构进行识别, 达到结构去噪的效果. 然而, 现有的工作无法有效在水质预测任务中对输入数据的时空特性与数值、结构噪声同时鲁棒建模. 例如, 不同污染物质由于水流复杂的流动模型产生沉降分层, 使得不同水质指标在观测节点间的拓扑关系可能存在较大的差异, 导致数值噪声与结构噪声的耦合关系更为复杂. 此外, 在各个观测节点间多维时序关系的拓扑关系可能各不相同, 利用现有方法建模数据的时序关系与拓扑关系容易在各个水质指标间互相掣肘, 难以有效同时处理两种噪声, 最终导致模型鲁棒性下降.

因此, 本文提出了 TaGNN 水质预测框架, 旨在通过利用水质历史数据中的趋势特征, 逐步解耦结构噪声与数值噪声, 从而解决水质预测中的时空数据噪声耦合问题. 该框架由 3 个模块组成: 趋势感知图结构学习模块, 用于构建多种潜在拓扑连接, 进而去除结构噪声; 去噪时空学习模块, 用于去除数值噪声; 节点对齐模块, 用于基于去噪后的多种拓扑连接聚合数值预测结果. 具体而言, TaGNN 首先利用历史水质数据随时间变化的趋势特征, 以原始结构信息为引导构建趋势图网络, 挖掘不同水质指标变化趋势间的相互关系, 初步排除原始数据中的数值噪声带来的负面影响; 随后通过多头分组潜在拓扑结构构建模块, 生成多种潜在可能的邻接矩阵, 有效解耦结构噪声. 在上述潜在真实可信的拓扑结构指导下, 去噪时空学习模块建立了多种潜在数值图网络, 深入挖掘各节点之间的指标数值时空依赖关系. 与此同时, 借助不确定性噪声预测模块, 实现了数值图网络节点表征中信息特征与噪声特征的有效解耦, 从而去除了数值噪声. 最终, 节点对齐模块通过评估不同趋势图网络与数值图网络中节点表征的对齐一致性, 计算对齐分数, 验证了各邻接矩阵的有效性, 聚合数值特征预测结果, 最终排除所有结构噪声, 作为最终水质预测的结果. 本文的主要贡献包括以下 3 个方面.

1) 提出一种基于水质指标趋势的图神经网络框架 TaGNN, 通过逐步解耦水质监测数据中的结构噪声与数值噪声, 实现了降噪水质预测.

2) 提出基于污染物沉降现象的多头结构提议-聚合方法, 捕捉监测节点水质指标间复杂的时空关系, 有效提取监测节点间的潜在拓扑关系。

3) 在某市某区真实水质监测数据集与额外加入人工数值噪声与结构噪声的增强数据集上, TaGNN 在 3 个不同指标上相较于 5 种基线算法均取得了优异的性能, 验证框架的有效性与鲁棒性。

1 相关工作

河道水质预测任务采集的数据主要包括 pH 值、溶解氧、氨氮等多元历史时序指标数据以及各个观测节点间的拓扑结构关系。因此, 本文首先调研现有的多元时空序列预测相关工作, 然后介绍与水质预测任务考虑时空数值/结构耦合噪声相关的带噪图神经网络研究进展。

1.1 多元时空序列预测

多元时间序列预测建模关键在于有效融合理解空间和时间相关性, 主要可以分为因子式和耦合神经网络架构两大类^[14-16]。

因子式的网络架构往往通过将时间特征与空间特征挖掘模块交替堆叠建模时空相关性。许多工作重点关注时空数据具有的多种类型时间特征提取, 再与空间特征结合进行建模。例如, 给定时间点信息可能与前一天甚至前一周同一时间的信息具有周期相似性。MLPST^[17]通过趋势性、周期性和一致性多粒度学习时间相关性, 空间相关性则体现在进行卷积的片(patch)以及片的特征单词(token)序列关系, 并基于全连接网络^[18]进行时空交互。鉴于在时空数据中普遍存在的短期和长期相关性, 探索多尺度时间相关性已成为改善时空学习的一个重要方向。GCN-M^[19]借助多尺度记忆模块学习流量数据中的局部时空特征和全局历史模式丰富流量嵌入用于标记动态流量状态, 通过结合静态节点嵌入和预定义的图信息生成动态图, 在此基础上与多个扩张卷积相结合, 以捕获流量嵌入中的时间和空间依赖性。MTMGNN^[20]通过近期和长期信息结合到一起, 以及 6 个包含静态和动态的空间相关图, 分别有效捕捉了时间和空间相关性。尽管这些结构可以有效地分别学习时间和空间相关性, 但对联合时空相关性建模能力不足, 使得跨越不同时间步长捕获复杂的时空关系具有挑战性。

另一方面, 耦合式架构则会综合挖掘数据的时空相关性, 相较于因子式架构更为简洁有效。ADGT-Net^[21]通过捕捉自适应双图表示模块的时间和空间特征, 分别作为 Transformer^[22]的编码和解码器, 有效捕捉了时空相关性。STAEformer^[23]引入了时空自适应嵌入模块, 发现时间序列相关性的同时对时序的位置信息进行捕获。STHAN^[24]通过采用具有多种边类型和元路径的时空异构图刻画区域之间的多种空间复合关系, 并在时空异构图中加入时间类型的边。WaveForm^[25]利用小波变换得到时域和频域上的时序特征, 学习全局图表示时序变量间关系。

已有一些工作成功将多元时间序列预测模型应用在河道水质监测。为了更好地处理河道水质数据在时间和空间上复杂依赖关系, 部分研究^[2,4]利用图卷积和 LSTM 分别捕获每个输入时刻监测点间的空间关系以及水质变化特征。宦娟等人^[1]通过构建时空图, 借助图卷积和 LSTM 实现京杭运河水质的预测。姜奕圻等人^[26]借助 GRU 实现了城市排水管网水质预测。刘萍等人^[27]借助 LSTM 实现了农村区域水环境的智能监测。翟爱丰等人^[28]提出了区间瞬时点源模型, 与水质水动力数学方程进行耦合, 有效应对三峡水库的突发污染事故。

1.2 带噪声的图神经网络训练

图神经网络 GNN 对于表达现实世界中复杂关系和结构具备天然优势, 在图数据建模分析方面表现出了卓越能力。然而实际应用中, 图数据常伴随节点标签和拓扑结构噪声。噪声信息很容易通过网络训练所依靠的图结构联通性和信息传递机制, 干扰其他节点或结构表征学习, 从而对模型性能和鲁棒性提出挑战。因此, 图神经网络的降噪方法逐渐受到关注。

图节点标签噪声治理往往利用半监督技术生成伪标签, 同时对可信与噪声节点连接的边结构分别做增加与删除, 从而保证真实信息被有效传播, 起到去噪效果。LAENNet^[29]将 GCN 预测的节点标签分布作为辅助特征, 与节点原始特征一起参与消息传递。LEGNN^[30]将估计的标签与给定真实标签相结合, 其中具有一致标签的边被赋予更大的权重。MTS-GNN^[31]则利用早期迭代中的模型来指导后续迭代中的模型训练、噪声标签去除和伪标签选择。

DropMessage^[32]通过随机删除消息矩阵中的元素,统一了所有基于消息传递的 GNN 中的随机删除策略.借助伪标签的引入,NRGNN^[10]将未标记节点与特征相似度高的标记节点连接起来,以减轻噪声标签的影响,并训练伪标签挖掘器选择伪标签.RTGNN^[11]在 NRGNN 基础上引入自强化和一致性正则化作为补充监督.STAL^[12]将主动学习和自我训练结合,在主动学习阶段利用不确定性策略选择最不确定的样本进行标记,然后结合自训练,使用模型生成高度确定的伪标签来扩展标签集.高图同态率可以减轻图标签噪声,LP4GLN^[33]通过图重构恢复同态性,利用标签传播校正噪声标签,并选择高置信度的标签进行节点分类.对比学习在处理标签噪声方面亦广受欢迎,DIOR^[34]通过双分区和对比学习,综合样本和参数两个层面,通过选择干净和噪声数据,引入半监督对比学习以促进一致性,设计关键和非关键参数的划分减轻噪声样本的记忆.CR-GNN^[35]通过无监督对比损失和邻域对比损失挖掘可靠节点.CGNN^[36]提出正则化项,通过图增强生成一致节点视图,抑制噪声节点过拟合,基于同质性样本选择技术识别并调整错误标签.异常检测与噪声过滤方面,GraphCleaner^[37]通过合成错误标签数据集,将错误标签作为负样本训练分类器,结合邻域感知错误标签检测,充分利用图的邻域依赖性,有效检测并纠正错误标签.

图结构噪声双向干扰输入拓扑结构和目标标签,导致模型性能下降及表征坍缩.为此,RGB^[38]通过信息瓶颈理论平衡图拓扑、图标签和图表征之间的相互依赖性,采用自监督学习和数据重参数化方法,成功应对双边噪声.SE-GSL^[13]结合结构熵和编码树理论,提供通用图结构学习框架,增强了图结构演化的可解释性.V-InFoR^[39]实现了处理轻微损坏时过滤冗余组件,同时识别和解释严重损坏,提高了 GNN 解释器在结构性损坏图上的性能.

然而,目前尚没有工作在水质预测这一时空任务中关注结构噪声与数值噪声的耦合问题,而多维时序数据的引入也大幅增加了两种噪声解耦的复杂性.因此,本文利用水质预测中的领域知识,挖掘多维时间序列的相互关系,逐步解耦并去除数值与结构噪声,对未来水质进行鲁棒预测.

2 基本定义

本节主要介绍水质预测问题的相关定义.

定义 1. 训练样本. 本文记训练输入的样本 $g = (X, A, y)$, 其中输入数据 $X \in \mathbb{R}^{N \times T \times C}$ 为当前关注的 N 个观测节点(后简称节点)提供的 T 个时间步中共 C 个水质指标数据,邻接矩阵 $A \in \{0, 1\}^{N \times N}$ 表示当前样本的原始邻接矩阵,即每个节点与其他节点的拓扑连接情况, $y \in \mathbb{R}^C$ 为真实的预测水质指标数据,即水质预测的回归目标. 本文假设在有限时间 T 内,水质预测相关水文拓扑结构不会随时间步产生明显变化.

定义 2. 趋势图网络. 本文定义每个样本 g 的趋势图网络为 $G' = (V, X', A)$, 其中 V 代表图中的节点集合, $\|V\| = N$; $X' \in \mathbb{R}^{N \times T \times C}$ 为对输入数据 X 在每一节点的每个指标上对时间维度 T 进行中心差分后得到的趋势图网络节点特征; A 为定义 1 中样本 g 的原始邻接矩阵,代表趋势图网络中节点的连接情况.

定义 3. 数值图网络. 定义每个样本 g 构建的数值图网络为 $G^p = (V, X, A^p)$. 同样的, V 代表图中节点集合; X 为 g 中原始输入数据; $A^p \in \{0, 1\}^{N \times N}$ 为数值图网络的邻接矩阵,代表原始的水文拓扑连接关系经过趋势图网络提取特征后,更新的节点拓扑连接关系.

问题 1. 水质预测问题. 本文定义水质预测问题的目标为使用 N 个节点提供 T 个时间步中 C 类水质指标的历史输入数据 g , 预测 $T+1$ 时刻中 C 类水质指标的真实数据. 假设目标的训练模型为 F , 预测得到的某一观测节点水质指标数据为 $y' \in \mathbb{R}^C$, 则问题可被定义为训练模型 F , 使得 $F(g) = y' \rightarrow y$.

3 TaGNN 框架

本文提出了一种面向噪声耦合水质预测的趋势感知图神经网络框架 TaGNN. 如图 1 所示, TaGNN 主要由趋势感知图结构学习模块,去噪时空学习模块以及节点对齐模块 3 部分组成. 框架首先输入原始数据每一节点中每个指标的中心差分数据 X' , 利用原始邻接构建趋势图网络 G' 矩阵, 通过趋势感知图结构学习模块构建得到多种潜在邻接矩阵拓扑结构 $\{A_i^p, i = 1, 2, \dots\}$, 初步降低原始数据中的结构噪声; 随后, 框架利用原始数据 X 与预测得到的 A_i^p 构建数值图网络 $\{G_i^p, i = 1, 2, \dots\}$ 提取节点特征对输入水质指标数值进行降噪, 最终通过趋势图网络 G' 与数值图网络 G_i^p 计算节点对齐分数, 聚合预测结果进行水质预测.

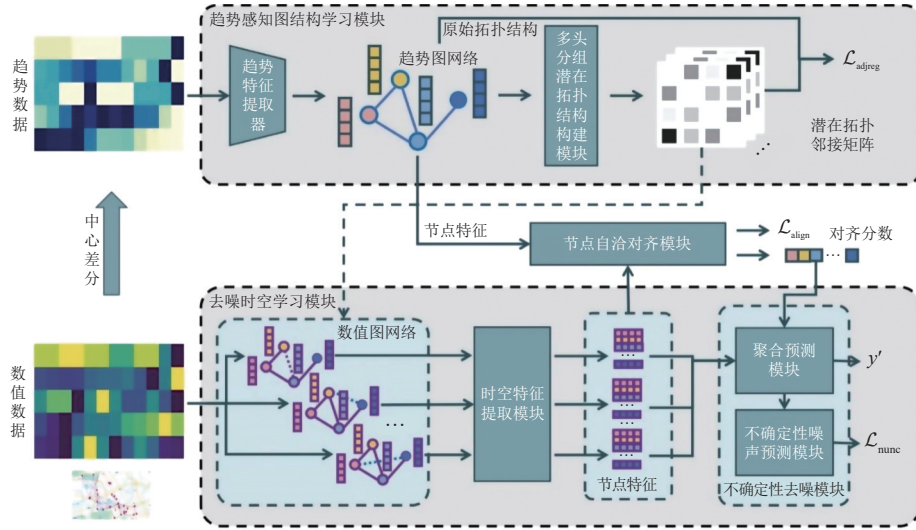


图1 TaGNN 框架流程示意图

3.1 趋势感知图结构学习模块

由于浮标传感器数值噪声的存在, 以及水质指标组间因为污染物的沉降与否和流向不同而产生的不同拓扑关系, 基于图神经网络的水质预测性能会急剧下降. 因此, 本文提出了一种趋势感知图结构学习模块, 利用传感器采集数值偏置的恒定性, 通过获取各个节点的趋势特征初步排除数值噪声, 同时利用节点指标组间趋势的相似性, 解耦并降低数据中的结构噪声, 减少错误信息的引入.

(1) 趋势特征提取

如图2(a)所示, 本文首先在通过中心差分得到趋势数据 X' 上, 利用卷积网络提取时间维度上的局部特征, 作为趋势图网络的初始化隐藏状态特征, 即:

$$\bar{h}_i^t = BN(\varphi(Conv(X'_i))) \quad (1)$$

其中, $\bar{h}_i^t \in \mathbb{R}^D$ 代表提取得到的第 i 个节点的趋势特征, $D \in \mathbb{R}$ 表示隐藏状态的特征维度, $X'_i \in \mathbb{R}^{C \times T}$ 表示 X' 中第 i 个节点的数据, $Conv(\cdot)$ 表示卷积网络, $\varphi(\cdot)$ 表示激活函数, $BN(\cdot)$ 表示批次标准化. 为保留不同节点之间的差异性, 本文也利用每个节点的编号对节点进行嵌入, 得到表征 $\hat{h}_i^t \in \mathbb{R}^D$. 随后, 将二者拼接经过全连接层得到每个节点在趋势图的隐藏表征 $\hat{h}_i^{(0)} \in \mathbb{R}^D$.

(2) 趋势图网络关系挖掘

为避免原始的带噪邻接矩阵 A 带来过多负面影响, 如图2(a)所示, 本文将原始邻接矩阵的频域信息作为引导, 利用 Transformer 挖掘节点间趋势特征的全局相互关系, 初步去除原始数据中包含的结构噪声. 受 Dwivedi 等人^[40] 的启发, 本文使用邻接拉普拉斯矩阵的特征向量作为位置编码, 即:

$$A^{Lap} = I - D^{-1/2} A D^{-1/2} \quad (2)$$

$$\lambda, EV = \text{eig}(A^{Lap}) \quad (3)$$

其中, $I \in \mathbb{R}^{N \times N}$ 表示单位矩阵, $D \in \mathbb{R}^{N \times N}$ 表示度矩阵, $\text{eig}(\cdot)$ 表示特征分解函数, λ 表示特征值, $EV \in \mathbb{R}^{N \times N}$ 表示分解得到的特征向量. 最终, 本文将位置编码 EV 经过一个全连接层升维后, 输入 Transformer 编码器挖掘节点趋势特征间的相互关系:

$$PE = MLP(EV^T) \quad (4)$$

$$h_i^{(0)} = \hat{h}_i^{(0)} + pe_i \quad (5)$$

$$H^{(l)} = \text{Transformer}(H^{(l-1)}) \quad (6)$$

其中, EV^T 表示特征向量 EV 的转置, $MLP(\cdot)$ 表示全连接层, $PE \in \mathbb{R}^{N \times D}$ 表示所有节点的位置编码矩阵, 而 $pe_i \in \mathbb{R}^D$ 则表示第 i 行对应节点的位置编码. $H^{(l)} \in \mathbb{R}^{N \times D}$ 表示第 l 层 Transformer 输出的所有节点隐藏状态拼接所得矩阵. 假设 Transformer 总共有 L 层, 本文将 $H^{(0)}$ 作为趋势图网络节点的最后表征, 记作 $H' = [h'_1, h'_2, \dots, h'_N]$.

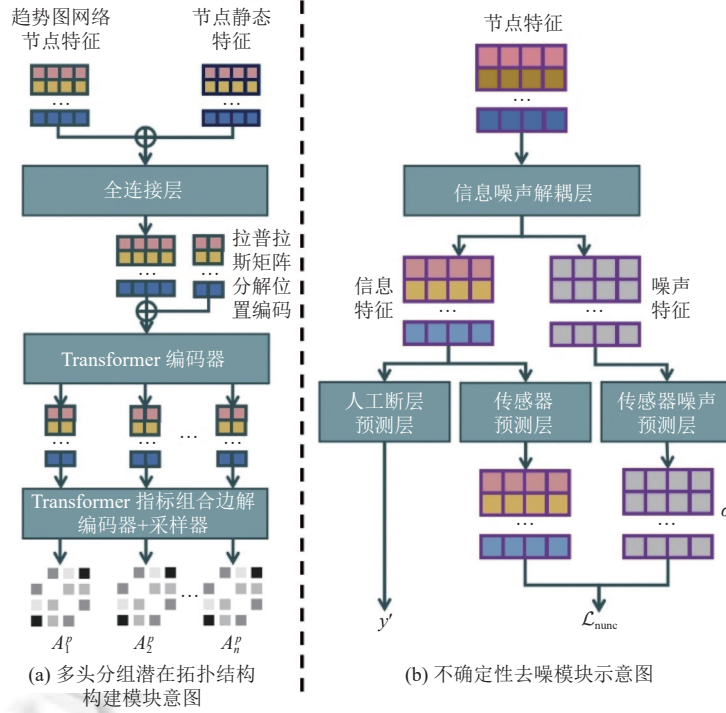


图2 TaGNN 框架子模块示意图

(3) 多头分组潜在拓扑结构构建

在经过 Transformer 编码器中的自注意力层和全连接层后, 得到的趋势图节点表征充分获取了节点之间的相互关系以及同一节点内各个指标之间的组合关系. 受 Transformer 的多头注意力机制的启发, 为了使得模型更好地关注指标的局部组合关系, 减少全局关系的互相掣肘, 本文将得到的趋势图网络节点表征 H' 拆分为 B 块趋势图网络节点表征 $S_b^{(0)} \in \mathbb{R}^{N \times D // B}$ ($b \in [1, B]$), 其中, $//$ 表示取整的除法. 随后, 这些表征被统一输入基于 Transformer 的 L 层边解码器, 进一步挖掘局部视角下各节点间的拓扑关系, 得到每一块的趋势图节点表征输出 $S_b^{(L)} \in \mathbb{R}^{N \times D // B}$, 记为 $S'_b = [s'_{1,b}, s'_{2,b}, \dots, s'_{N,b}]$.

为了获得更为鲁棒的邻接矩阵构建, 受 GTS^[41] 启发, 本文将邻接矩阵的每个元素视为独立的伯努利分布变量, 引入了 Gumbel 重参数化技巧对邻接矩阵进行重参数化, 即:

$$\theta_{i,j,b} = MLP(s'_{i,b} || s'_{j,b}) \quad (7)$$

$$g_{i,j,b}^1, g_{i,j,b}^2 \sim \text{Gumbel}(0, 1), i \in [1, N], j \in [1, N] \quad (8)$$

$$A_{i,j,b}^p = \text{Sigmoid}((\log(\theta_{i,j,b} / (1 - \theta_{i,j,b}))) + (g_{i,j,b}^1 - g_{i,j,b}^2) / s) \quad (9)$$

其中, $s \in [0, 1]$ 为可调节的温度, 当 $s \rightarrow 1$ 时, $A_{i,j,b}^p = 1$ 的概率为 $\theta_{i,j,b}$, $A_b^p \in \{0, 1\}^{N \times N}$ 为第 b 块预测得到的邻接矩阵, $A_{i,j,b}^p$ 则为 A_b^p 的第 i 行第 j 列的元素.

为了避免模型预测得到的邻接矩阵过于偏离真实情况, 引入额外的结构噪声. 本文加入了将重参数化的概率与真实的邻接矩阵 A 作交叉熵分类损失, 作为正则项引导模型构建的邻接矩阵符合真实拓扑结构, 即:

$$\mathcal{L}_{\text{adj}} = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \text{CELoss}(\theta_{i,j,b}, A_{i,j}) \quad (10)$$

3.2 去噪时空学习模块

在获得 B 种潜在的邻接矩阵 A_b^p 后, 本文将每个样本的输入数据 X 与对应的 A_b^p 结合, 构建 B 个数值图网络 G_b^p ($b \in [1, B]$). 随后, TaGNN 将构建的数值图网络 G_b^p 输入时空特征提取模块中提取原始数据的时空特征, 即数值图网络的节点表征 $H_b^p = [h_{1,b}^p, h_{2,b}^p, \dots, h_{N,b}^p] \in \mathbb{R}^{N \times D}$, 从而在通过趋势图网络去除结构噪声后进一步去除数值噪声. 其中时空特征提取模块并非本文的关注重点, 所有能够利用时序数据与邻接矩阵数据的模型都可以进行使用, 本文此处使用了 MTGNN 中给出的 GC 模块与 GT 模块交替挖掘数据中的时序特征与空间特征.

如图 2(b) 所示, 最后, TaGNN 使用不确定性去噪模块进行水质预测与自监督降噪. 当模型通过时空特征提取模块获得数值图网络的节点表征 H_b^p 后, 首先利用一个全连接层将 H_b^p 映射为两个相同形状的矩阵 $H_b^i \in \mathbb{R}^{N \times D}$ 以及 $H_b^n \in \mathbb{R}^{N \times D}$, 分别作为信息特征与噪声特征. 针对信息特征 H_b^i , 模型分别利用两个不同的全连接层, 分别预测每种邻接矩阵下的水质预测目标 $y_b' \in \mathbb{R}^C$ ($b \in [1, B]$) 以及针对最后一个时间步 T 的自监督预测 $Z_b \in \mathbb{R}^{N \times C}$.

另一方面, 模型使用另一个全连接层将噪声信息 H_b^n 预测为自监督预测结果 I_b 的偶然不确定性 $\sigma_b \in \mathbb{R}^{N \times C}$. 最后, 受 Kendall 等人^[42]的启发, 本文使用偶然不确定损失引导上述网络层解耦数值图网络节点表征中的信息特征与噪声特征, 即:

$$\mathcal{L}_{\text{unc}} = \frac{1}{N} \sum_{i=1}^N \frac{1}{2} \exp(-\log(\sigma_b^2)) (Z_b - X_T)^2 + \frac{1}{2} \log(\sigma_b^2) \quad (11)$$

其中, $\exp(\cdot)$ 表示自然指数函数, $X_T \in \mathbb{R}^{N \times C}$ 则表示原始输入数据 X 中所有节点在第 T 个时间步的数据.

3.3 节点对齐模块

在趋势感知图结构学习模块中, 模型利用趋势数据构建趋势图网络构建了 B 种潜在拓扑邻接矩阵 $\{A_b^p, b \in [1, B]\}$, 并结合数值数据构建数值图网络挖掘特征进行水质预测. 为确保进行有效的预测结果聚合, 需要估计各个构建的邻接矩阵下水质预测的效果. 由于趋势图网络中, 作为图节点信息趋势特征避免了传感器偏置导致的数值噪声引入, 原始邻接矩阵 A 中已有的边真实可信. 本文认为在趋势图网络预测所得的所有潜在拓扑结构中, 能够通过结合原始带噪数值数据, 还原已知水文拓扑结构以及水质指标趋势的潜在拓扑结构具有最高的合理性. 因此, 本文提出在构建的邻接矩阵下得到的数值图节点与趋势图中对应节点的对齐一致性能够表示各个构建所得的邻接矩阵的有效性. 在这一思路下, 本文提出利用验证趋势图网络 G' 与预测结构得到的数值图网络 G_b^p 节点表征的跨图节点对齐情况作为聚类依据的预测构建所的图一致性.

具体而言, 本文将趋势图网络 G' 中的节点表征, 与数值图网络 G_b^p 中对应节点的表征作为正样本, 并在其他所有节点中进行负采样得到的节点表征作为负样本, 本文记正样本对为 $(h_i^t, h_{i,b}^p)$, 负样本对为 $(h_i^t, h_{j,b}^p)$ ($i \neq j$). 受 Jiang 等人^[43]的启发, 本文使用类似于图数据中边预测的方法计算对齐分数, 即:

$$e_i^t = \text{MLP}_{\text{aligned}}(h_i^t) \quad (12)$$

$$e_{i,b}^p = \text{MLP}_{\text{aligned}}(h_{i,b}^p) \quad (13)$$

$$\text{score}_b = \frac{1}{N} \sum_{i=1}^N (e_i^t (e_{i,b}^p)^T - e_i^t (e_{j,b}^p)^T), i \neq j \quad (14)$$

针对第 3.2 节中 B 种构建的邻接矩阵下预测的 y_b^p , 本文利用 Softmax 函数对每一样本中 B 个 score_b 进行归一化后, 对预测结果进行聚合, 即:

$$\text{score}_b = \frac{\exp(\text{score}_b)}{\sum_{b=1}^B \exp(\text{score}_b)} \quad (15)$$

$$y' = \sum_{b=1}^B \text{score}_b \cdot y_b^p \quad (16)$$

3.4 训练损失

水质预测的训练目标是使得预测的 y' 接近于真实的水质指标 y , 因此本文采用平均平方误差 (mean square error, MSE) 作为损失函数, 即:

$$\mathcal{L}_{\text{reg}} = (y' - y)^2 \quad (17)$$

需要注意的是, 根据公式 (16), 损失函数 (公式 (17)) 可被改写为 $\mathcal{L}_{\text{reg}} = \sum_{b=1}^B \text{score}_b \cdot (y' - y)^2$. 在训练过程中如直接对 $\mathcal{L}_{\text{reg}}$ 进行优化可能导致 score_b 被动地倾向于在回归误差小的样本处上升而在回归误差大的样本处下降. 为了引导节点对齐模块的训练, 本文额外加入了对比学习损失, $\mathcal{L}_{\text{aligned}} = -\log(e'_i(e'_{i,b})^T - e'_i(e'_{j,b})^T)$, 并在计算回归损失 (公式 (17)) 时将 score_b 的梯度截断. 因此, 最终的损失函数为:

$$\mathcal{L} = \mathcal{L}_{\text{reg}} + \alpha \mathcal{L}_{\text{aligned}} + \beta \mathcal{L}_{\text{nunc}} + \gamma \mathcal{L}_{\text{adjreg}} \quad (18)$$

其中, α, β, γ 分别为预先设定的超参数.

4 实验

本文利用某市某区内监测到的真实河道水质数据集进行评估, 并与其他 5 种基准方法进行比较.

4.1 数据集和预处理

数据集记录了某段时期 38 个监测传感器收集的数据以及另一段期间 191 个人工断面收集的精确水质数据, 反映某市某区内部河道水质情况. 本文利用了水质检测最关注且空值最少的 pH、溶解氧、氨氮、氧化还原电位和透明度 5 个指标建模. 通过排除数据集过于稀疏的时间段, 并对剩余确实数据做插值操作后, 本文以每个人工断面的各个时间点为基准, 向前取 12 个时间步的传感器数据作为样本 g 构建数据集. 该数据集的主要信息如表 1 所示. 本文将水质监测传感器自动测得的数据作为输入数据, 将人工断面所获取的数据作为预测目标进行水质预测.

表 1 数据集的基本信息

观测节点	各样本节点数	各样本边数	时间点数	总数据量
传感器	38	108	26236	2922063
人工断面	1	2	95	7339

在数据预处理方面, 本文首先对数据集按照收集时间先后顺序排序后, 对数据集中的错误和缺失值进行了处理, 且对数据集中的空值进行了线性插值的填充. 为保证在同一时间点内尽可能出现较多的传感器观测值, 降低数据存在的稀疏性, 本文按照小时来划分时间点, 对每个时间点出现的传感器数目进行统计, 剔除掉传感器数目较少的时间点, 最终得到 26236 个时间点. 针对同一时间点内出现多次的传感器节点数据, 本文取平均值处理. 最后, 根据每一条人工断面预测目标所在时间点, 截取最邻近的 12 个时间点所对应的传感器监测数据作为历史数据输入模型预测.

为进一步验证 TaGNN 在水质预测场景中面对噪声问题的鲁棒性, 我们在采集的数据中分别额外引入了数值噪声和结构噪声. 具体而言, 我们在原始数据集上添加了随机扰动来模拟数值噪声, 并在真实邻接矩阵基础上随机去除了部分节点之间的连接, 以模拟结构噪声. 通过这种方式, 我们能够进一步评估模型在噪声干扰下的稳定性和鲁棒性.

4.2 基线方法和实验环境

本文使用了 5 种基准方法与所提模型进行性能对比测试, 包括处理时空数据的深度学习模型: MLPST、STAEformer、GTS、MTGNN; 以及扩散图卷积网络 DCRNN 与处理噪声数据的 Dropmessage 相结合的 DCRNN+Dropmessage.

1) MLPST^[17]. MLPST 借鉴 MLPmixer 的思想, 通过趋势性、周期性和一致性多粒度学习时间依赖性, 利用切片操作学习空间依赖性, 最终利用全连接网络进行时空交互.

2) STAEformer^[24]. STAEformer 引入了时空自适应嵌入模块, 在发现时间序列依赖性的同时对时间序列的位置信息进行了捕获.

3) GTS^[41]. GTS 通过学习概率图, 预测多个时间序列的离散图结构, 并与循环图卷积相结合来进行未来的时空预测.

4) MTGNN^[44]. MTGNN 设计了一个图学习模块来学习变量之间的空间单向相关性, 提出了具有混合跳跃传

播的图卷积和用于多变量时间序列预测的扩张初始层进行预测。

5) DCRNN^[45]+Dropmessage^[32]. DCRNN 利用图卷积网络捕获空间特征, 并应用基于 RNN 的编码器-解码器架构建模时间依赖性. Dropmessage 则是一种适用于消息传递 GNN 模型的通用随机删除方法, 通过对消息矩阵进行细粒度的删除操作, 允许一个节点将不同的消息传递给不同的邻居. 本文将 DCRNN 中的图卷积网络替换为 Dropmessage, 作为时空数据降噪工作的对比基线.

评估指标. 我们采用了 3 个广泛使用的评价指标, 包括平均绝对误差 (mean absolute error, MAE)、均方根误差 (rooted mean square error, RMSE) 和决定系数 (R^2) 评估模型有效性.

实验设置. 实验基于 PyTorch 和 PyG 实现了本文提出的模型与对比方法, 其中在 TaGNN 实现中, 设置 $B=4$, $\alpha, \beta, \gamma=0.1$. 设置实验使用 Adam 作为优化器, 批次大小设置为 64, 其余参数在实验中均按照各论文中提出的模型表现最佳参数配置. 实验采用容忍度为 30 的早停策略防止模型出现过拟合. 所有实验均使用 GeForce RTX 3090 GPU 进行. 为了确保实验结果真实公平的体现算法性能, 实验按照 3:1:1 将数据集划分为训练集、测试集和验证集进行 5 折交叉验证, 并将上述评估指标通过均值与标准差的形式进行展示. 实验结果如表 2 所示, 表中括号前数字表示指标均值, 括号内数字表示指标标准差, 加粗指标为最优模型指标, 带下划线指标为次优模型指标.

表 2 实验结果

模型	$R^2 \uparrow$	MAE (%) $\downarrow$	RMSE (%) $\downarrow$
MLPST	0.9864 (0.0013)	6.3654 (0.1393)	13.3098 (0.3562)
STAEformer	<u>0.9943 (0.0005)</u>	<u>5.1626 (0.0836)</u>	<u>10.3334 (0.4385)</u>
GTS	0.9463 (0.0051)	11.7682 (0.5959)	31.7703 (1.6017)
MTGNN	0.9936 (0.0007)	5.6423 (0.2469)	10.9385 (0.5930)
DCRNN+Dropmessage	0.8181 (0.0137)	20.5222 (1.0649)	58.4573 (1.6939)
TaGNN (本文)	0.9947 (0.0005)	4.9721 (0.1373)	9.9715 (0.4417)

4.3 实验分析

从表 2 和表 3 可得, 我们提出的 TaGNN 框架在水质预测任务中展现了较强的鲁棒性, 特别是在面对噪声耦合问题时. 具体而言, 表 3 的实验结果表明, 在同时存在数值噪声和结构噪声的情况下, TaGNN 对比目前的 SOTA 模型 STAEformer 在 R^2 上提升了 0.04%, 在 MAE 上提升了 4.32%, 在 RMSE 上提升了 3.71%; 对比与 TaGNN 结构类似的 GTS/MTGNN 中表现最好的模型 MTGNN 在 R^2 上提升了 0.06%, 在 MAE 上提升了 7.37%, 在 RMSE 上提升了 6.07%. 实验表明 TaGNN 在带有耦合数值噪声与结构噪声的水质预测问题中, 达到了目前最佳表现. 此外, 加入人工噪声后, 尽管 TaGNN 与大部分基线方法的均值表现有所下降, 但 TaGNN 的优势依然显著, 进一步验证了其在噪声处理方面的强大能力和鲁棒性.

表 3 实验结果 (含数值和结构噪声)

模型	$R^2 \uparrow$	MAE (%) $\downarrow$	RMSE (%) $\downarrow$
MLPST	0.9675 (0.0061)	10.0685 (0.8887)	24.6273 (2.4891)
STAEformer	<u>0.9942 (0.0006)</u>	<u>5.1668 (0.1315)</u>	<u>10.4433 (0.5153)</u>
GTS	0.9279 (0.0150)	13.4842 (1.8231)	36.6158 (3.7597)
MTGNN	0.9939 (0.0006)	5.3370 (0.1891)	10.7048 (0.4636)
DCRNN+Dropmessage	0.8904 (0.0148)	14.2838 (1.0014)	45.2871 (2.8365)
TaGNN (本文)	0.9946 (0.0006)	4.9435 (0.1289)	10.0555 (0.5661)

进一步分析, STAEformer 提出的时空自适应模块不同于其他因子式的时空模型, 耦合地建模了时间序列的时序关系以及时间序列间的位置关系, 得到了次优的结果, 而本文提出的 TaGNN 除了在噪声问题上考虑了结构噪声与数值噪声外, 同时针对水质预测问题不同指标中存在的结构噪声与数值噪声的耦合性, 对各个指标组合间的相互关系进行建模分析构建拓扑结构. 相对地, 在噪声处理方面, 当前的 SOTA 模型 STAEformer 以及 GTS、MTGNN 等只考虑了水质数据中的时间关系与空间关系, 没有针对数值噪声与结构噪声的耦合关系进行处理, 因

此表现不如本文提出的 TaGNN. 针对 DCRNN+Dropmessage, 虽然 DCRNN 能很好地提取水质数据的时空特征, Dropmessage 则通过对边的随机丢弃对数据进行数值降噪, 然而边的随机丢弃没有考虑到水质指标组间可能存在的不同的关系结构, 加剧了原始数据的结构噪声, 从而因为两种噪声的耦合问题进一步恶化了时序数据中的数值噪声, 降低模型对特征的提取能力. 因此 DCRNN+Dropmessage 在所有基线中表现最差, 这一结果也印证了本文在前文所述的不同指标组间存在不同的拓扑关系.

为了进一步验证所提方法的有效性, 我们对 5 种基线方法与 TaGNN 在 R^2 、MAE 和 RMSE 指标上的差异进行了配对样本 t 检验. 结果如表 4 和表 5 所示, 所有方法与 TaGNN 在 3 个指标上均存在显著差异 (P 值小于 0.05), 表明 TaGNN 相较于其他方法在这些指标上的差距具有统计学显著性, 特别是在与 STAEformer 的比较中, 尽管二者在均值上最为接近, 但我们的结果显示 TaGNN 在各项指标上均具有显著优势, 进一步验证了 TaGNN 的有效性.

表 4 差异显著性分析

模型	$P(R^2)$	$P(MAE)$	$P(RMSE)$
MLPST & TaGNN	0.0000***	0.0000***	0.0000***
STAEformer & TaGNN	0.0277*	0.0006***	0.0250*
GTS & TaGNN	0.0173**	0.0029**	0.0028**
MTGNN & TaGNN	0.0024**	0.0001***	0.0020**
DCRNN+Dropmessage & TaGNN	0.0000***	0.0000***	0.0000***

注: * $P<0.05$, ** $P<0.01$, *** $P<0.001$. 表格中数字表示 P 值, P 值小于 0.05, 拒绝原假设, 即数据之间存在显著差异, 下同

表 5 差异显著性分析 (含数值和结构噪声)

模型	$P(R^2)$	$P(MAE)$	$P(RMSE)$
MLPST & TaGNN	0.0009***	0.0004***	0.0003***
STAEformer & TaGNN	0.0457*	0.0007**	0.0415*
GTS & TaGNN	0.0008***	0.0007**	0.0001***
MTGNN & TaGNN	0.0049**	0.0021**	0.0054**
DCRNN+Dropmessage & TaGNN	0.0001***	0.0000***	0.0000***

4.4 消融实验

接下来, 我们通过消融实验证明所提 TaGNN 框架中各个模块的有效性. 其中, TaGNN (nunc-) 代表模型去掉了噪声不确定性模块; TaGNN (adjreg-) 代表模型去掉了邻接矩阵正则损失; TaGNN (ensemble-) 代表模型不再将图对齐结果作为聚合权重, 而简单的通过平均的方式进行聚合; TaGNN (nhead-) 代表模型不再使用多头注意力来提议多种潜在的拓扑结构, 而简单的通过单头自注意力预测一种降低结构噪声后的拓扑结构; TaGNN (---) 则代表模型仅包含 GTS 的图结构生成模块与 MTGNN 的时空提取模块. 实验结果如表 6 所示.

表 6 消融实验 (含数值和结构噪声)

模型	$R^2 \uparrow$	MAE (%) $\downarrow$	RMSE (%) $\downarrow$
TaGNN (nunc-)	0.9944 (0.0006)	4.9852 (0.1163)	10.2727 (0.5114)
TaGNN (adjreg-)	0.9920 (0.0056)	5.6754 (1.4797)	11.7344 (3.7379)
TaGNN (ensemble-)	0.9942 (0.0010)	5.0321 (0.2643)	10.4239 (0.8461)
TaGNN (nhead-)	0.9944 (0.0007)	4.9643 (0.1981)	10.1956 (0.6273)
TaGNN (---)	0.9941 (0.0010)	5.1132 (0.2556)	10.4956 (0.8434)
TaGNN (本文)	0.9946 (0.0006)	4.9435 (0.1289)	10.0555 (0.5661)

从表 6 中可得, TaGNN 中的各个模块均有效. 其中, 对比 TaGNN (nunc-) 与 TaGNN 的实验结果可得, 模型使用的噪声不确定性模块有效地解耦了数值噪声信息与真实信息, 为最终的预测提供了更鲁棒的特征信息; 此外, 通过对比 TaGNN (adjreg-) 与 TaGNN 的结果可得, 通过原始输入的邻接矩阵约束模型构建的邻接矩阵拓扑结构是有效的, 在训练初期引导模型构建拓扑结构的优化方向, 避免模型构建错误拓扑结构导致额外的结构噪声, 提高模型预测的准确性; 通过对比 TaGNN (ensemble-) 与 TaGNN 的结果可得, 模型使用的图对齐模块很好地建模了各种

潜在邻接矩阵拓扑结构的一致性,并根据 y 一致性对最终的多种预测结果进行聚合,得到了更为可靠的预测结果;而通过对比 TaGNN (nhead-) 与 TaGNN 的结果可得,通过提议-聚合模式来预测多种潜在的水文拓扑结构,并聚合这些拓扑下的预测结果具有更好的鲁棒性,从而证明了 TaGNN 中针对水质预测噪声耦合问题设计的提议-聚合机制的有效性;而 TaGNN (---) 的结果与其余实验结果对比可得,这 3 个模块的两两组合同样能起到较好的效果,避免相互掣肘,导致模型性能下降。

4.5 实际案例分析

最后,我们将模型预测得到的多种潜在可能的邻接矩阵进行了可视化,并与河道的真实情况进行对比验证,如图 3 所示,其中横纵坐标对应节点序号。如图 3(a)–(d) 所示,不同区块关注到不同指标组合间不同的拓扑结构。其中,如图 3(d) 所示,标记的深绿色所在列对应的节点为大量地下管网交汇处,因此与其余节点存在连接关系。此外,如图 3(e)、(f) 所示,图中高亮列所对应节点在实地勘察中发现人工排口,而模型预测得到的邻接矩阵在上述节点中对比原始邻接矩阵增加了新的边结构,与现实情况相符,证明 TaGNN 学习得到的邻接矩阵的有效性,即趋势图网络能够有效分离原始数据中带有结构噪声,为后续的数值图网络去噪提供有效支持。此外,通过线性层的参数分析,根据其权重数值大小,我们可以判断各个水质指标对不同预测邻接矩阵的贡献,从而方便相关机构发现尚未探明的人工排口或底下河道,使得模型也具有一定的可解释性。

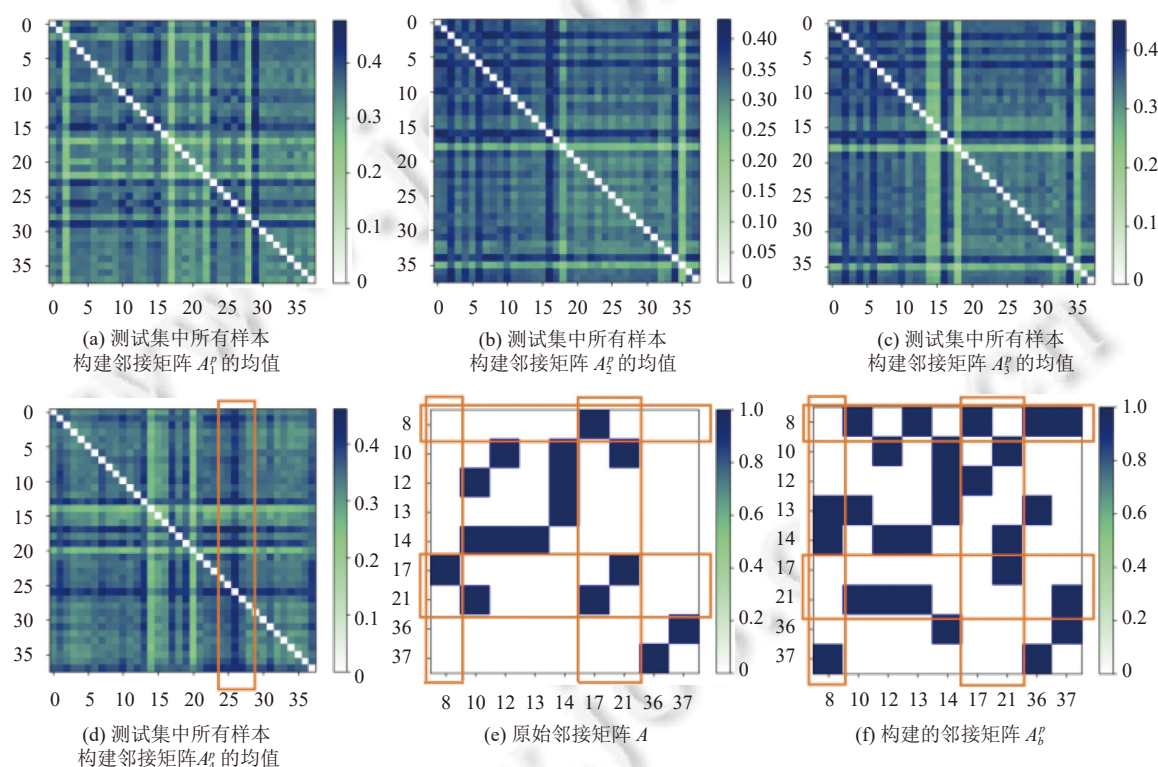


图 3 构建的拓扑结构可视化

5 总 结

本文提出了一种基于趋势感知图神经网络的耦合噪声水质预测框架 TaGNN, 该框架主要由趋势感知图结构学习模块利用趋势特征与原始拓扑结构构建趋势图网络, 建模指标间的组合关系, 预测观测节点间潜在的拓扑结构关系; 随后框架通过去噪时空学习模块利用构建的多种潜在邻接矩阵, 结合输入的原带噪声数据构建数值图网

络, 挖掘数据的时空特征, 并解耦噪声特征与信息特征, 最终通过信息特征进行预测. 此外, TaGNN 使用节点对齐模块, 对齐趋势图网络与数值图网络的对应节点特征, 作为聚合最终预测结果的依据. 通过在一个私有真实数据集上进行对比实验, 本文提出的 TaGNN 具有更高的 R^2 与更低的 MAE 与 RMSE. 因此本文的模型对于水质预测种耦合的结构噪声与数值噪声具有更好的鲁棒性. 未来的工作将专注于设计更适应水质预测任务的主干图神经网络, 进一步挖掘水质预测数据中的时空特征.

References:

- [1] Huan J, Zhang H, Xu XG, Yang BE, Shi B, Jiang JM. Research on the temporal and spatial prediction of the water quality of Beijing-Hangzhou canal based on graph convolution STG-LSTM. *China Rural Water and Hydropower*, 2022(8): 14–22 (in Chinese with English abstract).
- [2] Li YL, Zhang L, Li L. Study on water quality prediction of south source of Qiantang River based on GCN-LSTM. *Yellow River*, 2023, 45(12): 83–87, 95 (in Chinese with English abstract). [doi: [10.3969/j.issn.1000-1379.2023.12.015](https://doi.org/10.3969/j.issn.1000-1379.2023.12.015)]
- [3] Li XQ, Zheng H, Liu YY, Wan WH. Multi-source data machine learning-based study on method for regional water quality prediction. *Water Resources and Hydropower Engineering*, 2021, 52(11): 152–163 (in Chinese with English abstract). [doi: [10.13928/j.cnki.wrahe.2021.11.015](https://doi.org/10.13928/j.cnki.wrahe.2021.11.015)]
- [4] Xu JH, Wang JC, Chen L, Wu Y. Surface water quality prediction model based on graph neural network. *Journal of Zhejiang University (Engineering Science)*, 2021, 55(4): 601–607 (in Chinese with English abstract). [doi: [10.3785/j.issn.1008-973X.2021.04.001](https://doi.org/10.3785/j.issn.1008-973X.2021.04.001)]
- [5] Zhou J, Cui GQ, Hu SD, Zhang ZY, Yang C, Liu ZY, Wang LF, Li CC, Sun MS. Graph neural networks: A review of methods and applications. *AI Open*, 2020, 1: 57–81. [doi: [10.1016/j.aiopen.2021.01.001](https://doi.org/10.1016/j.aiopen.2021.01.001)]
- [6] Wu ZH, Pan SR, Chen FW, Long GD, Zhang CQ, Yu PS. A comprehensive survey on graph neural networks. *IEEE Trans. on Neural Networks and Learning Systems*, 2021, 32(1): 4–24. [doi: [10.1109/TNNLS.2020.2978386](https://doi.org/10.1109/TNNLS.2020.2978386)]
- [7] Bengio Y, Simard P, Frasconi P. Learning long-term dependencies with gradient descent is difficult. *IEEE Trans. on Neural Networks*, 1994, 5(2): 157–166. [doi: [10.1109/72.279181](https://doi.org/10.1109/72.279181)]
- [8] Gers FA, Schmidhuber J, Cummins F. Learning to forget: Continual prediction with LSTM. *Neural Computation*, 2000, 12(10): 2451–2471. [doi: [10.1162/089976600300015015](https://doi.org/10.1162/089976600300015015)]
- [9] Dey R, Salem FM. Gate-variants of gated recurrent unit (GRU) neural networks. In: *Proc. of the 60th Int'l Midwest Symp. on Circuits and Systems (MWSCAS)*. Boston: IEEE, 2017. 1597–1600. [doi: [10.1109/MWSCAS.2017.8053243](https://doi.org/10.1109/MWSCAS.2017.8053243)]
- [10] Dai EY, Aggarwal C, Wang SH. NRGNN: Learning a label noise resistant graph neural network on sparsely and noisily labeled graphs. In: *Proc. of the 27th ACM SIGKDD Conf. on Knowledge Discovery & Data Mining*. Singapore: ACM, 2021. 227–236. [doi: [10.1145/3447548.3467364](https://doi.org/10.1145/3447548.3467364)]
- [11] Qian SY, Ying HC, Hu RJ, Zhou JB, Chen JT, Chen DZ, Wu J. Robust training of graph neural networks via noise governance. In: *Proc. of the 16th ACM Int'l Conf. on Web Search and Data Mining*. Singapore: ACM, 2023. 607–615. [doi: [10.1145/3539597.3570369](https://doi.org/10.1145/3539597.3570369)]
- [12] Katsimpras G, Paliouras G. Improving graph neural networks by combining active learning with self-training. *Data Mining and Knowledge Discovery*, 2024, 38(1): 110–127. [doi: [10.1007/s10618-023-00959-z](https://doi.org/10.1007/s10618-023-00959-z)]
- [13] Zou DC, Peng H, Huang X, Yang RY, Li JX, Wu J, Liu CY, Yu PS. SE-GSL: A general and effective graph structure learning framework through structural entropy optimization. In: *Proc. of the 2023 ACM Web Conf*. Austin: ACM, 2023. 499–510. [doi: [10.1145/3543507.3583453](https://doi.org/10.1145/3543507.3583453)]
- [14] Jin GY, Liang YX, Fang YC, Shao ZZ, Huang JC, Zhang JB, Zheng Y. Spatio-temporal graph neural networks for predictive learning in urban computing: A survey. *IEEE Trans. on Knowledge and Data Engineering*, 2024, 36(10): 5388–5408. [doi: [10.1109/TKDE.2023.3333824](https://doi.org/10.1109/TKDE.2023.3333824)]
- [15] Guo SN, Lin YF, Feng N, Song C, Wan HY. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting. In: *Proc. of the 33rd AAAI Conf. on Artificial Intelligence*. Honolulu: AAAI, 2019. 922–929. [doi: [10.1609/aaai.v33i01.3301922](https://doi.org/10.1609/aaai.v33i01.3301922)]
- [16] Jiang JW, Han CK, Zhao WX, Wang JY. PDFormer: Propagation delay-aware dynamic long-range Transformer for traffic flow prediction. In: *Proc. of the 37th AAAI Conf. on Artificial Intelligence*. Washington: AAAI, 2023. 4365–4373. [doi: [10.1609/aaai.v37i4.25556](https://doi.org/10.1609/aaai.v37i4.25556)]
- [17] Zhang ZJ, Huang Z, Hu ZW, Zhao XY, Wang WY, Liu ZT, Zhang JB, Qin SJ, Zhao HW. MLPST: MLP is all you need for spatio-temporal prediction. In: *Proc. of the 32nd ACM Int'l Conf. on Information and Knowledge Management*. Birmingham: ACM, 2023. 3381–3390. [doi: [10.1145/3583780.3614969](https://doi.org/10.1145/3583780.3614969)]
- [18] Tolstikhin IO, Housby N, Kolesnikov A, *et al.* MLP-Mixer: An all-MLP architecture for vision. In: *Proc. of 35th Int'l Conf. on Neural Information Processing Systems*. Curran Associates Inc., 2021. 1857.

- [19] Zuo JW, Zeitouni K, Taher Y, Garcia-Rodriguez S. Graph convolutional networks for traffic forecasting with missing values. *Data Mining and Knowledge Discovery*, 2023, 37(2): 913–947. [doi: [10.1007/s10618-022-00903-7](https://doi.org/10.1007/s10618-022-00903-7)]
- [20] Yin D, Jiang RH, Deng JW, Li YK, Xie Y, Wang ZY, Zhou YF, Song X, Shang JS. MTMGNN: Multi-time multi-graph neural network for metro passenger flow prediction. *GeoInformatica*, 2023, 27(1): 77–105. [doi: [10.1007/s10707-022-00466-1](https://doi.org/10.1007/s10707-022-00466-1)]
- [21] Kong JL, Fan XM, Zuo M, Deveci M, Jin XB, Zhong KY. ADCT-Net: Adaptive traffic forecasting neural network via dual-graphic cross-fused Transformer. *Information Fusion*, 2024, 103: 102122. [doi: [10.1016/j.inffus.2023.102122](https://doi.org/10.1016/j.inffus.2023.102122)]
- [22] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I. Attention is all you need. In: *Proc. of the 31st Int'l Conf. on Neural Information Processing Systems*. Long Beach: Curran Associates Inc., 2017. 6000–6010.
- [23] Liu HC, Dong Z, Jiang RH, Deng JW, Deng JL, Chen QJ, Song X. Spatio-temporal adaptive embedding makes vanilla Transformer SOTA for traffic forecasting. In: *Proc. of the 32nd ACM Int'l Conf. on Information and Knowledge Management*. Birmingham: ACM, 2023. 4125–4129. [doi: [10.1145/3583780.3615160](https://doi.org/10.1145/3583780.3615160)]
- [24] Ling S, Yu Z, Cao SS, Zhang HP, Hu S. STHAN: Transportation demand forecasting with compound spatio-temporal relationships. *ACM Trans. on Knowledge Discovery from Data*, 2023, 17(4): 54. [doi: [10.1145/3565578](https://doi.org/10.1145/3565578)]
- [25] Yang FH, Li X, Wang M, Zang HY, Pang W, Wang MZ. WaveForM: Graph enhanced wavelet learning for long sequence forecasting of multivariate time series. In: *Proc. of the 37th AAAI Conf. on Artificial Intelligence*. Washington: AAAI, 2023. 10754–10761. [doi: [10.1609/aaai.v37i9.26276](https://doi.org/10.1609/aaai.v37i9.26276)]
- [26] Jiang YQ. Research on water quality prediction model of urban sewer networks based on multi-source data fusion [Ph.D. Thesis]. Harbin: Harbin Institute of Technology, 2022 (in Chinese with English abstract). [doi: [10.27061/d.cnki.ghgdu.2022.000338](https://doi.org/10.27061/d.cnki.ghgdu.2022.000338)]
- [27] Liu P. Study of smart monitoring and prediction method of rural water environment based on Internet of Things [Ph.D. Thesis]. Yangzhou: Yangzhou University, 2020 (in Chinese with English abstract). [doi: [10.27441/d.cnki.gyzdu.2020.002521](https://doi.org/10.27441/d.cnki.gyzdu.2020.002521)]
- [28] Zhai AF. Research on Three Gorges Reservoir operation scheme based on water quality safety guarantee of drinking water source [Ph.D. Thesis]. Beijing: North China Electric Power University (Beijing), 2022 (in Chinese with English abstract). [doi: [10.27140/d.cnki.ghbbu.2022.000029](https://doi.org/10.27140/d.cnki.ghbbu.2022.000029)]
- [29] Zhang CX, Li XM, Pei HB, Zhang ZJ, Liu B, Yang B. LaenNet: Learning robust GCNs by propagating labels. *Neural Networks*, 2023, 168: 652–664. [doi: [10.1016/j.neunet.2023.09.035](https://doi.org/10.1016/j.neunet.2023.09.035)]
- [30] Guo K, Cao XF, Liu ZN, Chang Y. Taming over-smoothing representation on heterophilic graphs. *Information Sciences*, 2023, 647: 119463. [doi: [10.1016/j.ins.2023.119463](https://doi.org/10.1016/j.ins.2023.119463)]
- [31] Liu YN, Wu ZQ, Lu ZY, Wen GQ, Ma JB, Lu GQ, Zhu XF. Multi-teacher self-training for semi-supervised node classification with noisy labels. In: *Proc. of the 31st ACM Int'l Conf. on Multimedia*. Ottawa: ACM, 2023. 2946–2954. [doi: [10.1145/3581783.3613117](https://doi.org/10.1145/3581783.3613117)]
- [32] Fang TR, Xiao ZQ, Wang CP, Xu JR, Yang X, Yang Y. DropMessage: Unifying random dropping for graph neural networks. *Proc. of the 37th AAAI Conf. on Artificial Intelligence*. Washington: AAAI, 2023. 4267–4275. [doi: [10.1609/aaai.v37i4.25545](https://doi.org/10.1609/aaai.v37i4.25545)]
- [33] Cheng Y, Shan CH, Shen YF, Li X, Luo SQ, Li DS. Label propagation for graph label noise. *arXiv:2310.16560*, 2023.
- [34] Wang HX, Jiang HY, Sun JN, Zhang SK, Chen C, Hua XS, Luo X. DIOR: Learning to hash with label noise via dual partition and contrastive learning. *IEEE Trans. on Knowledge and Data Engineering*, 2024, 36(4): 1502–1517. [doi: [10.1109/TKDE.2023.3312109](https://doi.org/10.1109/TKDE.2023.3312109)]
- [35] Li XX, Li QY, Li D, Qian HD, Wang JY. Contrastive learning of graphs under label noise. *Neural Networks*, 2024, 172: 106113. [doi: [10.1016/j.neunet.2024.106113](https://doi.org/10.1016/j.neunet.2024.106113)]
- [36] Yuan JY, Luo X, Qin YF, Zhang YS, Ju W, Zhang M. Learning on graphs under label noise. In: *Proc. of the 2023 IEEE Int'l Conf. on Acoustics, Speech and Signal Processing (ICASSP)*. Rhodes Island: IEEE, 2023. 1–5. [doi: [10.1109/ICASSP49357.2023.10096088](https://doi.org/10.1109/ICASSP49357.2023.10096088)]
- [37] Li YW, Miao X, Hooi B. GraphCleaner: Detecting mislabelled samples in popular graph learning benchmarks. In: *Proc. of the 40th Int'l Conf. on Machine Learning*. Honolulu: JMLR.org, 2023. 832.
- [38] Zhou ZK, Yao JC, Liu JX, Guo XW, Yao QM, He L, Wang L, Zheng B, Han B. Combating bilateral edge noise for robust link prediction. In: *Proc. of the 37th Int'l Conf. on Neural Information Processing Systems*. New Orleans: Curran Associates Inc., 2023. 934.
- [39] Wang SZ, Yin J, Li CZ, Xie X, Wang JX. V-InFoR: A robust graph neural networks explainer for structurally corrupted graphs. In: *Proc. of the 37th Int'l Conf. on Neural Information Processing Systems*. New Orleans: Curran Associates Inc., 2023. 2464.
- [40] Dwivedi VP, Bresson X. A generalization of Transformer networks to graphs. *arXiv:2012.09699*, 2020.
- [41] Shang C, Chen J, Bi JB. Discrete graph structure learning for forecasting multiple time series. In: *Proc. of the 9th Int'l Conf. on Learning Representations*. OpenReview.net, 2021.
- [42] Kendall A, Gal Y. What uncertainties do we need in Bayesian deep learning for computer vision? In: *Proc. of the 31st Int'l Conf. on Neural Information Processing Systems*. Long Beach: Curran Associates Inc., 2017. 5580–5590.
- [43] Jiang M. Cross-network learning with partially aligned graph convolutional networks. In: *Proc. of the 27th ACM SIGKDD Conf. on Knowledge Discovery & Data Mining*. Singapore: ACM, 2021. 746–755. [doi: [10.1145/3447548.3467282](https://doi.org/10.1145/3447548.3467282)]

- [44] Wu ZH, Pan SR, Long GD, Jiang J, Chang XJ, Zhang CQ. Connecting the dots: Multivariate time series forecasting with graph neural networks. In: Proc. of the 26th ACM SIGKDD Int'l Conf. on Knowledge Discovery & Data Mining. New York: ACM, 2020. 753–763. [doi: 10.1145/3394486.3403118]
- [45] Li YG, Yu R, Shahabi C, Liu Y. Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. In: Proc. of the 6th Int'l Conf. on Learning Representations. Vancouver: OpenReview.net, 2018. 678–685.

附中文参考文献:

- [1] 宦娟, 张浩, 徐宪根, 杨贝尔, 史兵, 蒋建明. 基于图卷积 STG-LSTM 的京杭运河水质时空预测研究. 中国农村水利水电, 2022(8): 14–22.
- [2] 李余隆, 张兰, 李立. 基于 GCN-LSTM 的钱塘江南源水质预测研究. 人民黄河, 2023, 45(12): 83–87, 95. [doi: 10.3969/j.issn.1000-1379.2023.12.015]
- [3] 李雪清, 郑航, 刘悦忆, 万文华. 基于多源数据机器学习的区域水质预测方法研究. 水利水电技术 (中英文), 2021, 52(11): 152–163. [doi: 10.13928/j.cnki.wrahe.2021.11.015]
- [4] 许佳辉, 王敬昌, 陈岭, 吴勇. 基于图神经网络的地表水水质预测模型. 浙江大学学报 (工学版), 2021, 55(4): 601–607. [doi: 10.3785/j.issn.1008-973X.2021.04.001]
- [26] 姜奕圻. 基于多源数据融合的城市排水管网水质预测模型研究 [博士学位论文]. 哈尔滨: 哈尔滨工业大学, 2022. [doi: 10.27061/d.cnki.ghgdu.2022.000338]
- [27] 刘萍. 基于物联网的农村区域水环境智能监测及预测方法研究 [博士学位论文]. 扬州: 扬州大学, 2020. [doi: 10.27441/d.cnki.gyzdu.2020.002521]
- [28] 翟爱丰. 基于饮用水源地水质安全保障的三峡水库调度方案研究 [博士学位论文]. 北京: 华北电力大学 (北京), 2022. [doi: 10.27140/d.cnki.ghbbu.2022.000029]



孙建明(1977—), 男, 高级工程师, 主要研究领域为城市运行管理, 生态环境领域大数据应用.



张啸(1991—), 男, 博士, 副教授, CCF 专业会员, 主要研究领域为数据挖掘, 分布式学习, 联邦学习.



徐宇扬(1999—), 男, 博士生, 主要研究领域为数据挖掘, 图神经网络, 大模型以及相关应用.



庄福振(1983—), 男, 博士, 教授, CCF 高级会员, 主要研究领域为迁移学习, 机器学习, 数据挖掘, 推荐系统.



仝硕(2001—), 男, 硕士生, 主要研究领域为医学领域的数据挖掘.



吴健(1975—), 男, 博士, 教授, CCF 杰出会员, 主要研究领域为医学人工智能.



应豪超(1991—), 男, 博士, 副研究员, CCF 专业会员, 主要研究领域为数据挖掘, 图神经网络应用.