

面向金融风险预测的时序图神经网络综述*

宋凌云, 马卓源, 李战怀, 尚学群

(西北工业大学 计算机学院, 陕西 西安 710072)

通信作者: 宋凌云, E-mail: lysong@nwpu.edu.cn; 尚学群, E-mail: shang@nwpu.edu.cn



摘要: 金融风险预测在金融市场监管和金融投资中扮演重要角色, 近年来已成为人工智能和金融科技领域的热门研究主题. 由于金融事件的实体之间存在复杂的投资、供应等关系, 现有的金融风险预测研究常利用各种静态和动态的图结构来建模金融实体间的关系, 并通过卷积图神经网络等方法将相关的图结构信息嵌入金融实体的特征表示中, 使其能够同时表征金融风险相关的语义和结构信息. 然而, 以前的金融风险预测综述仅关注了基于静态图结构的研究, 这些研究忽视了金融事件中实体间关系会随时间动态变化的特性, 降低了风险预测结果的准确性. 随着时序图神经网络的发展, 越来越多的研究开始关注基于动态图结构的金融风险预测, 对这些研究进行系统、全面的回顾有助于学习者构建面向金融风险预测研究的完整认知. 根据从动态图中提取时序信息的不同途径, 首先综述 3 类不同的时序图神经网络模型. 然后, 根据不同的图学习任务, 分类介绍股价趋势风险预测, 贷款违约风险预测, 欺诈交易风险预测, 以及洗钱和逃税风险预测共 4 个领域的金融风险预测研究. 最后, 总结现有时序图神经网络模型在金融风险预测方面遇到的难题和挑战, 并展望未来研究的潜在方向.

关键词: 时序图神经网络; 金融风险预测; 股价趋势风险; 贷款违约风险; 欺诈交易风险; 洗钱和逃税风险

中图法分类号: TP18

中文引用格式: 宋凌云, 马卓源, 李战怀, 尚学群. 面向金融风险预测的时序图神经网络综述. 软件学报, 2024, 35(8): 3897–3922. <http://www.jos.org.cn/1000-9825/7087.htm>

英文引用格式: Song LY, Ma ZY, Li ZH, Shang XQ. Review on Temporal Graph Neural Networks for Financial Risk Prediction. Ruan Jian Xue Bao/Journal of Software, 2024, 35(8): 3897–3922 (in Chinese). <http://www.jos.org.cn/1000-9825/7087.htm>

Review on Temporal Graph Neural Networks for Financial Risk Prediction

SONG Ling-Yun, MA Zhuo-Yuan, LI Zhan-Huai, SHANG Xue-Qun

(School of Computer Science and Technology, Northwestern Polytechnic University, Xi'an 710072, China)

Abstract: Financial risk prediction plays an important role in financial market regulation and financial investment, and has become a research hotspot in artificial intelligence and financial technology in recent years. Due to the complex investment, supply and other relationships among financial event entities, existing research on financial risk prediction often employs various static and dynamic graph structures to model the relationship among financial entities. Meanwhile, convolutional graph neural networks and other methods are adopted to embed relevant graph structure information into the feature representation of financial entities, which enables the representation of both semantic and structural information related to financial risks. However, previous reviews of financial risk prediction only focus on studies based on static graph structures, but ignore the characteristics that the relationship among entities in financial events will change dynamically over time, which reduces the accuracy of risk prediction results. With the development of temporal graph neural networks, increasingly more studies have begun to pay attention to financial risk prediction based on dynamic graph structures, and a systematic and comprehensive review of these studies will help learners foster a complete understanding of financial risk prediction research. According to different methods to extract temporal information from dynamic graphs, this study first reviews three different neural network models for

* 基金项目: 国家重点研发计划 (2020AAA0108504); 国家自然科学基金 (62102321); 中央高校基本科研业务费专项资金 (D5000230095); 陕西省重点研发计划 (2021ZDLGY03-08)
收稿时间: 2023-02-20; 修改时间: 2023-06-16; 采用时间: 2023-11-15; jos 在线出版时间: 2024-03-20
CNKI 网络首发时间: 2024-03-23

temporal graphs. Then, based on different graph learning tasks, it introduces the research on financial risk prediction in four areas, including stock price trend risk prediction, loan default risk prediction, fraud transaction risk prediction, and money laundering and tax evasion risk prediction. Finally, the difficulties and challenges facing the existing temporal graph neural network models in financial risk prediction are summarized, and potential directions for future research are prospected.

Key words: temporal graph neural network (TGNN); financial risk prediction; stock price trend risk; loan default risk; transaction fraud risk; money laundering and tax evasion risk

随着互联网技术的进步,近年来金融交易活动同时在线上和线下迎来了巨大的发展,涌现出了许多新的金融活动,例如在线保险、网上购物、网络贷款等.金融市场繁荣发展的同时,也给金融风险的预测带来了巨大的挑战和困难.金融风险(例如,欺诈交易和贷款违约风险)严重威胁着金融市场的繁荣与稳定.如果对金融风险预测不准确或者不及时,可能会引发系统性的全球金融危机(如2008年美国次贷危机),这将对经济的发展和人民生活水平产生严重的负面影响.为此,近年来有大量的工作致力于金融风险预测研究.

传统的金融风险预测研究大多采用统计机器学习的方法,它们试图利用风险相关的金融指标训练金融风险分类器.虽然能够取得一定的效果,但这些方法有两个局限性:1)忽视了金融实体之间的复杂关系,这些关系对金融风险的预测有重要作用^[1-3];2)不适用金融指标数据缺失的金融实体,这使得它们无法应对大量中小企业的风险预测^[4-6].为了解决上述问题,近年来研究人员利用各种图结构来建模金融风险事件中金融实体间的复杂关系,提出了大量面向金融风险预测的图神经网络,它们大致可以被分为两类:基于静态图的图卷积神经网络和基于动态图的时序图神经网络(temporal graph neural network, TGNN).前者虽然能够建模金融风险中实体间的复杂关系,但却忽视了金融风险事件的关键因素,即时间信息.为了解决这个问题,研究人员利用允许节点和边随时间发生变化的时序图(temporal graph)(在本文中,时序图和动态图是等同的(exchangeable))来建模金融实体间的关系,提出了一系列的面向金融风险预测的时序图神经网络,其能够有效捕捉金融实体间关系的时序特征,提升风险预测结果的准确性.

近年来,神经网络(neural network, NN)和图神经网络(graph neural network, GNN)的发展推动了多个任务的进步,比如图像分类^[7-9],目标检测^[10-13],视觉问答^[14-17],跨媒体检索^[18-22],虚假新闻检测^[23-25],异构信息网络^[26-30]嵌入等.已有研究综述了基于静态图的图卷积网络在金融风险预测领域的进展^[31-34].例如,Wang等人^[31]介绍了图神经网络在金融领域的应用,Kurshan等人^[32,33]介绍了基于图计算的金融犯罪和欺诈检测,Rajput等人^[34]提出时序图学习在金融领域应用的简要教程.然而,纵观现有文献,还未有对时序图神经网络在常见金融风险预测领域全面、详细的综述.为此,本文分析和总结了在金融风险预测领域中最新的时序图神经网络研究,并讨论了此领域中现存的挑战和未来的研究方向.

具体地,本文的核心贡献可以总结如下.

1)据我们所知,本文是第1个系统地总结时序图神经网络模型在金融风险预测领域的综述,其总结了金融数据建模常用时序图的类型和结构,并根据对时序图中时间信息不同的建模方式,对面向金融风险预测的时序图神经网络模型进行了归纳和总结.

2)本文提出了一种基于时序图神经网络的金融风险预测研究的综述框架,其按照图学习任务类别对相关研究进行了统一的划分,并总结讨论了基于不同图学习任务的风险预测方法的数据来源和时空信息处理方式.

3)本文总结了当前用于金融风险预测的研究在数据集获取和构建、时空特征建模、小样本学习、模型计算效率和可解释性等5个方面的难题和挑战,并对未来金融领域风险预测的研究方向作出了展望.

本文第1节介绍了时序图的定义和分类.第2节根据时间信息的不同建模方法,将时序图神经网络模型粗略地划分为以下3个类别:基于循环神经网络(recurrent neural network, RNN)的、基于卷积神经网络(convolutional neural network, CNN)的和基于注意力机制(attention mechanism)的方法.第3节详细介绍和分析了时序图神经网络模型在4大金融风险预测子领域的应用,包括股票趋势风险预测,贷款违约风险预测,欺诈交易风险预测,以及洗钱和逃税风险预测.在介绍每个风险预测子领域的方法时,本文按照不同的图学习任务对这些方法进行了具体的划分和介绍,同时对它们的建模方式和其所用数据进行了总结.第4节总结和展望了时序图神经网络在金融风

险预测领域的挑战和潜在研究方向. 最后对全文进行了总结.

1 时序图

本节首先介绍了时序图的基本定义和惯用符号表示, 然后按照图的建模方式对时序图进行了分类和概述, 具体介绍内容如下.

1.1 时序图定义

时序图是含有时间信息的普通图, 其节点和边可能会随时间发生变化, 例如节点或边的增加和减少, 节点的属性和的边权值的改变. 最大时间步长为 T 的所有时序图的集合表示为 $G = \{g_t\}_{t=1}^T$, 其中, t 时刻的图表示为 $g_t = (V_t, E_t, A_t)$, V_t 和 E_t 是图 g_t 节点和边的集合, A_t 是节点的邻接矩阵. $X(V_t)$ 是图 g_t 中所有节点的特征矩阵, $x(v_i^t)$ 是图 g_t 中节点 v_i^t 的特征表示, $X(E_t)$ 是图 g_t 中所有边的特征矩阵, $x(e_{i,j}^t)$ 是边 $e_{i,j}^t$ 的特征表示, H_t 是图 g_t 经过卷积模块隐藏层的输出.

1.2 时序图分类

现有研究^[35,36]通常利用如下两类图结构来构建和表示时序图, 即基于图快照 (graph snapshot) 的离散时序图和基于时间戳 (timestamp) 的连续时序图.

1.2.1 离散时序图

离散时序图由图快照按时间顺序排列而成, 每个图快照是一个普通图, 表示时序图在某时刻的状态. 现有的时序图神经网络大多利用传统的图神经网络 (GNN) 编码单个图快照的空间结构信息, 同时利用循环神经网络 (RNN) 编码时序图的时间序列信息, 从而完成对离散时序图的处理. 然而, 将动态图表示为离散时序图存在以下问题: 1) 单个图快照的时间信息被忽略. 2) 粗粒度的离散时间序列不足以精确表示连续的时序信息. 3) 时间步长的选择不能兼顾计算效率和精度: 紧凑的时间步长使得离散时间更加逼近连续时间, 但由于图快照数量的增加导致计算效率不高; 虽然增加时间步长可以提升计算效率, 但是会导致更大的预测精度损失. 根据账户交易记录生成的离散时序图如图 1 所示, 3 张图快照按时间序列排列, 已存在的交易用绿色线段标注, 新产生的交易用黄色线段标注. 图中账户之间仅具有空间关系, 用蓝色虚线表示, 不同时刻的账户之间具有时间关系, 用橙色虚线表示.

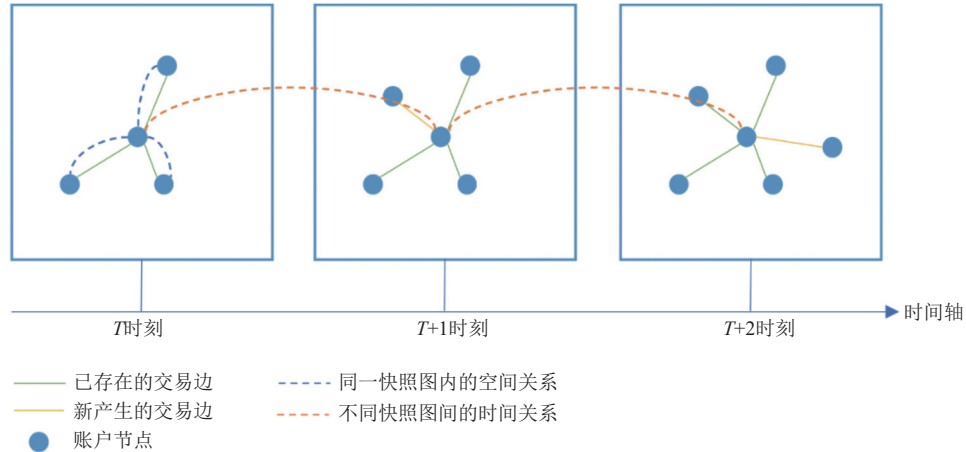


图 1 基于图快照的离散时序图

1.2.2 连续时序图

连续时序图中节点间的边附带有时间戳, 可以记录时间段或者某时刻的信息, 通常被用来建模具体事件, 可记为四元组 $(h, r, t, \Delta\tau)$. 其中, h , r , t 和 $\Delta\tau$ 分别表示头实体节点, 关联信息, 尾实体节点和事件的持续时间. 其中, 事件的持续时间 $\Delta\tau$ 可以简化为时刻 τ , 此时四元组重写为 (h, r, t, τ) , 其表示某事件瞬时发生后即消失, 比如两个联系

人之间发送一次邮件. 基于前文描述, 若将连续时间信息简化为时刻信息, 或是按某特定时间间隔对基于时间戳的连续时序图数据采样, 则基于时间戳的连续时序图将退化成基于图快照的离散时序图. 连续时序图的样例如图 2 所示, 图中的实体可分为独立账户和经推断属于同一实际控制人或集团的账户集合. 例如, 独立实体 A 和 D 的关系可用四元组 $(A, r_{AD}, E, \Delta\tau_{AD})$ 表示, 实体集合 E 和独立实体 A 的关系可用四元式 $(A, r_{AE}, E, \Delta\tau_{AE})$ 表示.

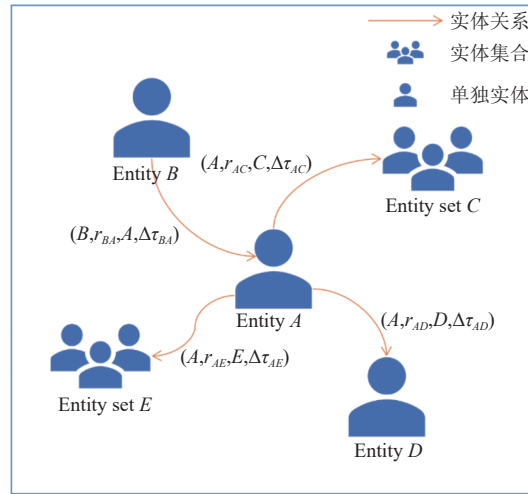


图 2 基于时间戳的连续时序图

1.2.3 面向金融领域风险事件的图构建

连续时序图蕴含的时间信息更丰富, 能够最大程度保留真实的时间信息. 然而, 由于要处理连续时间信息, 连续时序图的模型搭建比离散时序图复杂, 不能简单地通过堆叠 GNN 和 RNN 架构实现. 因此, 连续时序图更适合短时间内有较大发展和变化的事件建模, 而离散时序图适合较长时间内情况相对稳定的事件的建模^[35,36].

据我们所知, 金融领域的事件在大多数情况下均采用了离散时序图的建模方式. 例如, 公司间相互持股的情况比较稳定, 即使增加新的持股关系通常也能维持较长时间. 因此, 公司的股权关系变化适合用离散时序图建模. 在贷款担保交易中, 贷款从批准借出到归还完本金、利息的过程需要持续一段时间, 故贷款担保事件也适合用离散时序图表示. 因此, 研究人员应当首先根据所研究问题(事件)所持续的时间长短确定其所采用的时序图建模方式.

当前研究采用离散图的形式构建金融领域图数据的方式大致可分为两类, 一类是遵照数据(如股票价格数据)本身的时序属性, 按照固定时间间隔采样具有时序属性的数据对象, 再根据数据对象间的关系(如股票-行业类型关系和持股关系等)生成每个时刻图快照中的边, 从而构建基于特定时间间隔的离散时序图; 二是基于金融事件(如比特币交易和银行卡转账)的时间信息, 给不同金融实体间关系附上时间属性, 从而生成由具有时间属性的边构成的离散时序图.

如上所述, 金融事件如公司交叉持股和金融借贷通常在短时间内难以有较大变化故适合用离散时序图来对其建模; 此外, 股市的风险预测往往依赖于对股票群体中长期的联动变化趋势的建模, 其同样适用于离散时序图. 近年来, 有部分学者开展了连续时序图的表示和学习研究^[37-42], 但其具体实验大多采用 Wikipedia (<http://snap.stanford.edu/jodie/wikipedia.csv>), Reddit (<http://snap.stanford.edu/jodie/reddit.csv>) 等数据集. 这些数据集中数据对象间的边变化频繁, 与金融风险预测领域中的数据有较大差异. 目前暂未发现有构建连续时序图进行金融风险预测的研究.

2 时序图神经网络模型

由于金融领域风险事件通常更适合用离散时序图建模, 故本文仅关注基于离散时序图的时序图神经网络模型. 时序图神经网络能够同时建模不同图快照序列中节点之间的时间依赖关系和每个图快照中节点之间的空间或语义等关系, 并已在风险预测^[43-47]、交通预测^[48-54]和人类动作识别^[55-58]等领域有着广泛的应用. 针对时序图中节

点间的空间和语义等关联的建模,现有研究常利用图嵌入^[59,60]或以图卷积网络 (graph convolutional network, GCN)^[61]为代表的图神经网络等方法。为了捕捉不同时刻图快照之间的时间依赖关系,现有的时序神经网络模型可以大致分为以下3类:基于循环神经网络 (RNN) 的、基于卷积神经网络 (CNN) 的,以及基于注意力机制 (attention mechanism)^[62]的模型。

在具体介绍各个方法之前,首先结合4大金融风险子领域,对上述3类时序图神经网络模型进行分类总结和概述,具体信息如表1所示。

表1 基于时序图神经网络的金融风险预测模型

金融风险预测领域	处理时序数据方法	具体方法名称	模型名称
股票变动风险预测	RNN方法	LSTM GRU	RSR, TRAN, ST-Trader, Price graphs, HAD-GNN HGTAN
	CNN方法	时间维度卷积 GLU	KDTCN, GraphWaveNet, MTGNN, StemGNN, STHGCN TEDHGN
	注意力方法	Transformer Hawkes process	THGNN STHAN-SR
贷款违约风险预测	RNN方法	LSTM GRU	DDPF, TDEP, ST-GNN, TemGNN TRACER, DGANN
	RNN方法	LSTM	DyHGN-DE
欺诈交易风险预测	CNN方法	时间维度卷积	DHGReg, STAGN
	注意力方法	Transformer	DyHGN-HGT, APAN
	其他方法	LambdaNetworks 时序点过程	BRIGHT TeGraF
洗钱和逃税风险预测	RNN方法	LSTM RWA	MGC-LSTM, Temporal-GCN T-EGAT
	其他方法	直接拼接时间嵌入	DELATOR

(1) 采用RNN方法提取时间信息的TGNN: 股价变动预测子领域的RSR^[63], TRAN^[64], ST-Trader^[65], Price graphs^[66], HAD-GNN^[67]和贷款违约预测子领域的DDPF^[68], TDEP^[69], ST-GNN^[70], TemGNN^[46], 欺诈交易风险预测的DyHGN-DE^[71], 以及洗钱风险预测的MGC-LSTM^[72], Temporal-GCN^[73]均采用LSTM^[74]处理时间序列。股价变动预测子领域的HGTAN^[75]和贷款违约风险预测的TRACER^[76]和DGANN^[77]使用GRU^[78]处理时间序列, 比LSTM训练效率更高。逃税风险预测的T-EGAT^[79]采用类似RNN架构的循环加权平均单位 (recurrent weighted average unit, RWA)^[80]来处理时间信息。

(2) 采用CNN方法提取时间信息的TGNN: 股价变动预测子领域KDTCN^[81]和3个时空图神经网络 (spatial-temporal graph neural network) (时空图神经网络为特殊的时序图神经网络, 指针对时空图进行建模的图神经网络。时空图指特殊的时序图, 指节点和边不随时间发生增减、仅权重等数值发生变化的时序图。常用于交通流量预测和人体动作识别场景) 模型GraphWaveNet^[82], MTGNN^[83], StemGNN^[84], 以及超图模型STHGCN^[85]采用基于时序卷积来并行处理时间信息, TEDHGN^[86]采用门控线性单元 (gated linear unit, GLU)^[87]进一步提高计算效率。欺诈交易风险预测的DHGReg^[88]基于普通1-D CNN架构处理时间信息, STAGN^[89]采用3D卷积处理含有时间、空间信息的三维张量。

(3) 采用注意力方法提取时间信息的TGNN: 比如在股价变动预测子领域THGNN^[90]使用Transformer^[91]对时间序列编码; 欺诈交易风险预测领域的DyHGN-HGT^[71]和APAN^[92]使用Transformer架构来建模时间信息。

(4) 其他TGNN: 比如在股价变动预测子领域的STHAN-SR^[93]使用Hawkes process^[94]来进一步增强时间注意力, 欺诈交易风险预测领域的BRIGHT^[95]和TeGraF^[96]分别使用LambdaNetworks^[97]和时序点过程 (temporal point process, TPP)^[98]来建模时间信息。逃税风险预测的DELATOR^[99]通过简单地拼接时间嵌入来建模时间信息。除利用RNN、CNN和注意力机制以外, 采用其他TGNN方法处理时序数据的现有研究较少, 尚未形成确定的处理流

程范式, 因此本节后续内容并未单独列出该分类。

2.1 基于 RNN 方法提取时间信息

根据提取空间、时间信息的先后顺序, 这类方法可被分为以下两组: 第 1 组方法, 首先通过使用图卷积编码每个时刻图快照的结构信息, 然后再利用 LSTM, GRU 等 RNN 模型来建模不同时刻图快照间的时间依赖关系。比如 Hou 等人^[65]提出的预测股票价格走势的时空深度神经网络, 其首先利用分钟级的股票价格数据, 将公司间隐藏的依赖关系建模为图。然后, 利用 GCN-LSTM 框架, 将股票之间隐藏关系的空间信号纳入到股票价格的时间序列预测中。这种先利用 GNN 建模空间信息, 然后再用 RNN 处理时序图的建模方法也常被用于基于用户交易记录图的金融欺诈风险预测任务^[71]。

第 2 组方法, 先用 RNN 等模型来处理包含时间信息的序列数据, 再利用 GNN 方法聚合信息以捕捉图的空间结构特征。例如, Ying 等人^[64]提出的用于股票排名预测的时间感知图关系注意力网络 (time-aware graph relational attention network, TRAN)。首先, 基于行业分类构建股票关系图, 分别从股票价格历史序列和股票描述文档中提取股票的历史特征和股票属性; 然后, 设计时间感知图关系注意力网络, 通过股票的历史特征与股票属性的动态交互获取股票间的关系强度, 并将其视为股票关系图上的注意力权重。最后, 利用图卷积将股票关系图的结构特征嵌入到股票的特征表达中, 将得到的股票特征用于排名预测。

综上, 建模时序图中的时空信息时, 最常见是采用基于 RNN 与图卷积结合的方法。然而, RNN 类方法处理时间信息存在迭代传播大量耗时和梯度爆炸、梯度消失问题, 且不适合对长距离的时间序列进行建模。

2.2 基于 CNN 方法提取时间信息

这类方法使用卷积操作来处理时序数据, 能够以非递归的方式并行处理时序图, 具有并行计算、梯度稳定和内存需求低的优势。例如, Deng 等人^[81]提出了面向股票价格趋势预测的知识驱动的时序卷积网络 (temporal convolutional network, TCN)^[100-102], 其首先利用外部知识图谱学习财经新闻事件的图嵌入表示, 得到空间特征。然后, 通过拼接股票价格的时间序列信息和前文提到的图嵌入, 实现时间和空间信息的整合, 并将拼接后的结果送入到 TCN 中进行处理。TCN 的卷积层采用因果卷积, 使得某时刻的输出仅能通过卷积该时刻之前的结果得到; 采用扩张卷积, 使得模型适合长序列时序数据的预测; 每层均由相同长度的 1D 全卷积网络 (1D fully-convolutional network, FCN) 架构组成, 保证隐藏层和输出结果与输入序列保持相同长度。TCN 类算法的核心架构示意图, 如图 3 所示。

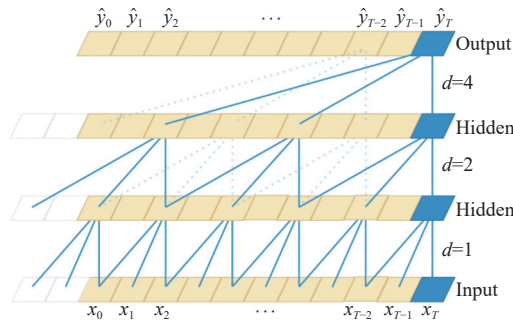


图 3 TCN 架构的示意图

2.3 基于注意力机制提取时间信息

这类方法通常将时间序列的建模方法和 GNN 集成到一个统一的网络架构中, 并采用注意力机制来抽取相应的时序特征。一个代表性的方法是 Xiang 等人^[90]提出的用于股票价格变动预测的时序异构图神经网络模型。如图 4 所示, 该模型首先生成不同时刻的股票关系图快照, 进行正负邻居节点采样, 基于位置编码为不同时刻的节点作以区分, 然后用自注意力机制 (self-attention mechanism) 处理时序信息进行历史价格编码。然后采用时序图多头注意力机制 (multi-head attention mechanism) 区分不同实体节点的重要性, 获得图的结构信息, 最后采用异构图神经网络

将该时序图的空间结构信息嵌入节点的特征表达. 该方法使用了完全基于注意力机制的 Transformer 模型, 其本质是采用注意力机制提取时间信息. 该模型的优点是能够有效捕获时序数据中的长距离时间依赖关系, 且计算效率高, 但是 Transformer 包含的众多矩阵运算会带来内存占用过大的问题.

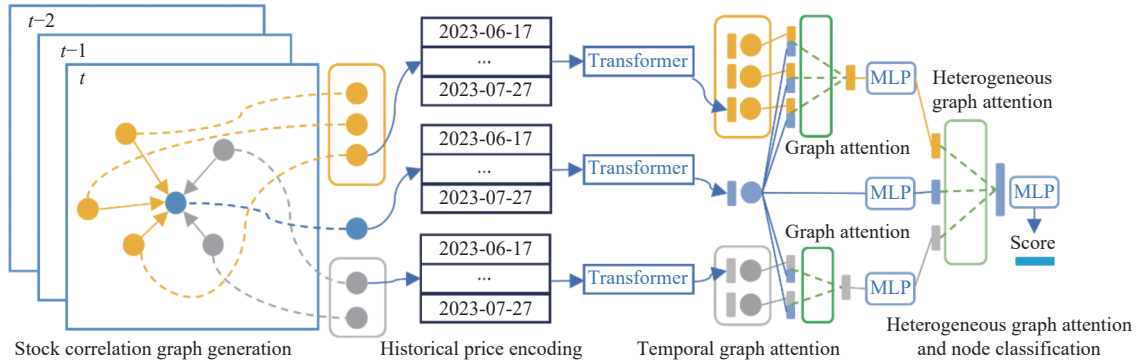


图 4 基于注意力机制提取时序特征的流程图

3 基于时序图神经网络的金融风险预测

现有基于时序图神经网络的金融风险预测研究主要关注股价趋势预测, 欺诈交易预测, 贷款违约预测, 以及洗钱和逃税风险预测等 4 个方面的问题. 在已有研究中, 上述金融风险预测任务通常被建模为节点分类、节点回归、边分类和链接预测等图学习任务, 其目的是预测未来时刻时序图中节点或边的属性值. 具体地, 各任务的形式化定义如下.

(1) 节点分类任务: 时序图可表示为 $G(g_1, g_2, \dots, g_T)$, 其中 $g_t = (V_t, A_t, X_t)$ ($1 \leq t \leq T$) 表示第 t 个时间步的图快照, V_t 表示第 t 时间步所有节点的集合, A_t 表示该时间步图快照的邻接矩阵, X_t 是该时间步的图中节点的特征矩阵. 给定时序图集合 G 和有标记的节点的集合 V_t , 时序图中节点分类任务的目的是对标签未知的节点 $v \in V_u$ 进行分类. 股票价格变化趋势的涨跌风险预测, 潜在欺诈用户分类的欺诈交易风险, 潜在不合规用户识别的洗钱和逃税风险预测等任务, 都可建模为图中节点是否异常的二分类问题.

(2) 节点回归任务: 给定时序图 $G(g_1, g_2, \dots, g_T)$, 未来某时刻图中金融实体节点的风险数值预测可看作是单步或多步的节点时间序列回归问题. 具体地, 单步或多步节点时间序列回归任务是基于时序图中实体节点的历史金融数据, 预测实体节点在未来某个时刻或多个时刻和风险相关的属性值的任务. 以股票价格变动风险预测为例, 设 $\mathbf{z}_t \in \mathbb{R}^N$ 表示时刻 t 的 N 维向量, 该向量可以是多家公司在当前时刻 t 的股票价格或风险数值等. 给定历史序列 $X = \{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_t\}$, 节点时间序列回归任务是在历史观察的基础上预测节点在未来单个或多个时刻的值, 即预测序列 $Y = \{\hat{\mathbf{z}}_{t+1}, \hat{\mathbf{z}}_{t+2}, \dots, \hat{\mathbf{z}}_{t+H}\}$, 其中 H 表示需要预测的时刻数目, 称为预测视界. 当预测视界 $H > 1$ 时, 即为多步回归预测.

(3) 边分类任务: 给定时序图 $G(g_1, g_2, \dots, g_T)$, 在 t 时刻 $g_t = (V_t, E_t)$ ($1 \leq t \leq T$). 对于图中每个边 $e'_{i,j} \in E_t$, 以贷款违约风险预测为例, 目标是生成一个可学习的风险评分函数 $f(e'_{i,j})$. 风险评分值表示边 $e'_{i,j}$ 的异常度, 更高的风险评分 $f(e'_{i,j})$ 代表着当前借贷关系 $(v'_i, e'_{i,j})$ 有更高的风险. 若对异常评分 $f(e'_{i,j})$ 设置阈值, 将高于阈值的异常得分设为有风险, 低于阈值的异常得分视为正常, 可将基于时序图的贷款违约风险预测建模为边的二分类问题.

(4) 链接预测任务: 给定时序图 $G(g_1, g_2, \dots, g_T)$ 和图中有标签的边集合 $E = \{(u, v, y)\}$ ($1 \leq t \leq T$), 其中 $y = 1$ 表示节点 u 和 v 之间存在边, $y = 0$ 表示它们之间不存在边. 以基于供应链挖掘的风险预测研究为例, 其中供应链的挖掘可建模为基于时序图的链接预测任务, 目标是用于确定两个公司之间是否存在供应链关系, 并将供应链关系的状态引入金融风险的预测过程中.

(5) 排名预测任务: 需要说明的是, 与传统方法将股票价格变动风险预测建模为节点分类问题或回归问题不同, 此任务的目的是学习一个股票收益排名函数 $\hat{r}^{t+1} = f(X^t)$, 其中 X^t 是 N 支股票在 t 个交易日内的价格数据. 通过 $f(X^t)$, 将一系列股票映射到一个排名列表. 股票的排名越高, 则其在时刻 $t+1$ 处的预期投资收益越高, 风险越小.

3.1 股价趋势风险预测

股价趋势风险是指未来发生了不乐观的经济活动, 或由于各种事先无法预料的消极因素的影响, 造成股票实际收益比预期收益发生减少的可能性. 作出相对准确的股票价格下跌预测能在一定程度上规避资产的损失风险, 预测股票价格上涨或是挑选出预测盈利能力最好的某些股票, 甚至可以逆转资产损失的风险为资产增值的机遇. 因此, 股票价格预测与股票风险预测息息相关.

最早的股票价格预测使用统计学方法如整合移动平均自回归模型 (autoregressive integrated moving average model, ARIMA)^[103]和广义自回归条件异方差模型 (general autoregressive conditional heteroskedasticity model, GARCH)^[104]等. 近年来, 基于支持向量机 (support vector machine, SVM)^[105], 随机森林 (random forest)^[106]等机器学习方法和基于长短期记忆网络 (long short-term memory, LSTM)^[74], 门控循环单元 (gated recurrent unit, GRU)^[78]等深度学习方法也被广泛用于基于股票价格序列的风险分析. 此类研究均基于一个基本假设, 认为各支股票之间是彼此独立的, 但它们只考虑了价格信息. 根据有效市场假说^[107], 目标公司股价的变化会受到与之相关的各公司股价变动的影 响, 不同股票之间存在难以仅从股价数据反映出来的复杂联系, 比如是否属于同一行业, 股东增、减持情况, 相互持股关系等信息. 这些复杂的显式或隐式的属性或关联信息需要用图结构去捕捉. 此外, 股票价格本身是时间序列数据, 把握其时间演化特征也很重要, 故传统方法和基于静态图的图神经网络方法不再适用.

综上所述, 采用时序图神经网络预测股价趋势从而规避股票投资风险, 是近年来的研究热点, 也是本节的主要内容. 现有基于时序图神经网络的方法通常基于股票价格的时序图, 将股价风险预测问题建模为图中节点分类、节点回归或排名预测任务, 具体介绍如下.

3.1.1 节点分类

这类方法将股票趋势风险预测任务建模为基于股价时序图的节点分类任务, 其中节点表示公司股票, 预测目标是股价的变动趋势 y . 该任务主要预测股票价格的二元变动趋势, 其中 $y = 1$ 表示股价的上升趋势, $y = 0$ 表示股价的下降趋势. 具体地, y 的定义可表示为:

$$y = \begin{cases} 1, & \text{if } x_t > x_{t-1} \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

其中, x_t 表示交易日 t 的股票价格. 例如, Deng 等人^[81]提出了面向股票趋势突变预测的知识驱动时序卷积网络 KDTCN, 如图 5 所示, 从非结构化的新闻文本中抽取事件三元组构成知识图谱, 然后将知识图谱嵌入特征, 与价格向量拼接好后送入 TCN 模型 (如图 3) 进行分类. 该方法将知识驱动的事件编码引入到了股票趋势的二分类任务中, 解决了现有方法对股票趋势突变不够敏感和预测结果不可解释的问题.

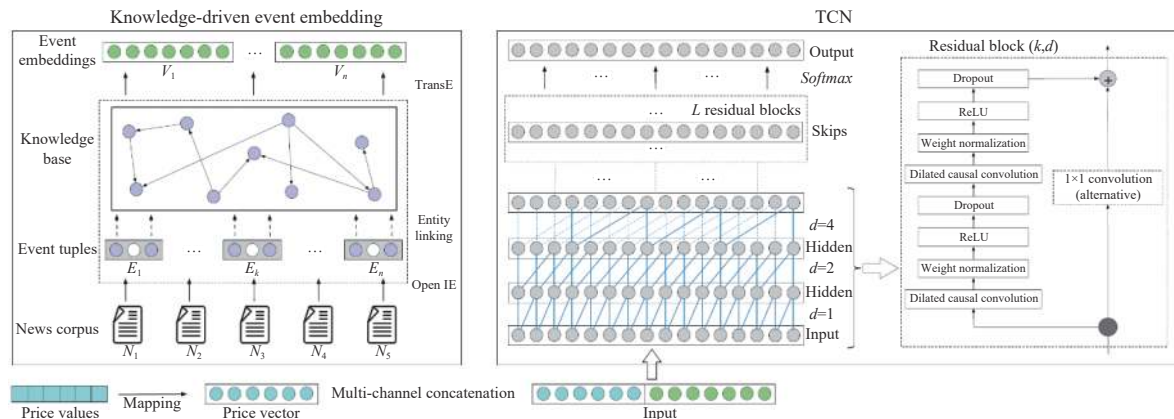


图 5 KDTCN 架构示意图

然而, 上述方法难以有效建模股票序列数据中的长期依赖关系和股票价格的非平稳突变特点. 为此, Wu 等人^[66]提出如图 5 的算法流程, 该算法由时间序列嵌入模块和预测模块组成. 第 1 个模块首先使用可视化图算法^[108]将股

票价格序列映射为图, 然后采用 struc2vec^[109]学习能够保持不同时间点股价之间关联结构的节点嵌入, 并采用集体影响 (collective influence, CI) 算法^[110]测量价格图的节点权重. 第 2 个模块是基于神经网络的预测模块, 其为了弥补股票时间序列结构信息中时间序列的丢失, 采用基于注意力机制的 RNN 来解决不同时间点股票价格的隐式长期依赖和非平稳突变问题. 然后, 利用自注意力层建模股票间的关联结构. 最后, 利用所得到的每个股票的隐藏表示, 预测相应股票价格的涨跌变化趋势.

Wu 等人^[66]提出的上述方法没有区分不同股票的历史价格对目标股票价格趋势预测的重要性. 为此, 研究人员提出了基于股票动态图的时序图神经网络. 例如, Tian 等人^[67]提出了混合注意力动态图神经网络 HAD-GNN, 其首先根据不同股票历史价格之间的相似性构建了两种类型的动态股票联动图, 包含节点分类和回归预测: 1) 用于股票运动方向预测的曼哈顿股票联动图, 即分类任务. 2) 用于股票价格预测任务的皮尔逊股票联动图, 即回归任务. 然后, HAD-GNN 采用编码器-解码器架构进行时空信息处理. 其中, 混合注意力编码器使用基于 LSTM 的时序注意力层和节点级注意力机制来捕获股票时序图中历史股价的时间和空间特征, 而作为解码器的预测模块负责输出股票变动趋势和回报率的预测结果.

以上方法在构建图时, 利用从股票描述文档和财经新闻等文本中提取股票属性和投资关系等信息, 存在关系提取不精确的问题, 并且引入的外部知识图谱难以反映公司间关系的动态变化. 为此, Xiang 等人^[90]提出了面向股价趋势预测的时序异构图神经网络模型. 与传统的基于知识图谱的关系或通过人类标注构建关系不同, 该模型首先基于真实的股票历史市场信号^[111,112]之间的相关系数来得到相关矩阵, 而不需要额外的领域知识或者新闻文本. 然后, 根据相关矩阵中各元素的值确定公司之间正向的 (相关性>阈值) 或负向的 (相关性<阈值) 的关系, 初步构建异构图. 随后, 利用 Transformer 进行股票的时序价格编码, 用两阶段时序图注意力机制聚合来自正、负关系的股价信息, 并用异构图注意力机制更新股票节点的特征表达. 最后, 将股票的特征送入分类器进行股票价格趋势变动的节点二分类预测.

然而, 上述方法均是利用股票间的成对关系构建股票价格时序图. 目前, 许多研究者基于超图 (hypergraph) 提出了对于属于同一类别股票的共同变动趋势预测新方法. 超图中的超边 (hyperedge) 可以连接多个节点, 一条超边包含的多个节点之间具有高阶关系 (high-order relation). 针对传统两两关系建模方法包含大量噪声的问题, Sawhney 等人^[85]考虑对股票所属行业间关系进行建模, 提出了时空超图卷积神经网络 STHGCN 来建模股票价格的时间演变特征和属于相同行业股票之间的高阶关系. 具体而言, 如图 6 所示, STHGCN 首先根据板块和行业信息构建股票超图, 在构造的超图上进行空间超图卷积以提取空间信息. 然后, 基于 GRU 使用门控时间卷积来学习股票特征的时间演化特征. 最后, 将门控的时间超图卷积和空间超图卷积组合为一个时空超图卷积块, 将股票之间的相关性信息和股价的时间演化信息嵌入到股票节点的特征表达中. 最后, 将所学股票特征送入分类器进行股价趋势风险的节点分类预测.

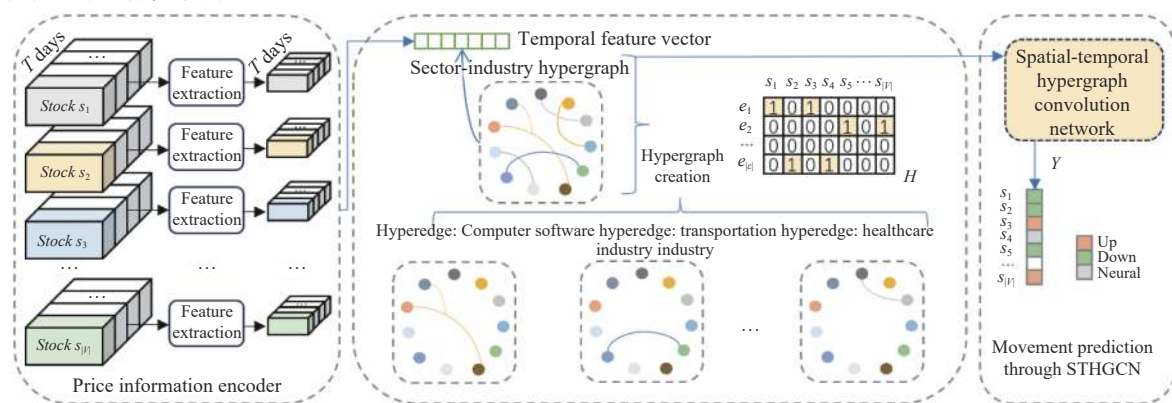


图 6 时空超图卷积网络 STHGCN 的模型图

针对 STHGCN 只能平等地对待不同权重超边的问题, Cui 等人^[75]提出了一种时间感知的超图多重注意力网络 HGTAN, 其同时考虑了超边内 (intra-hyperedge)、超边间 (inter-hyperedge) 和超图间 (interhypergraph) 这 3 类注

意力, 来分别衡量不同节点、超边和超图的重要性, 有效地区分了不同结构信息的重要性.

3.1.2 节点回归

股票价格或投资回报率等数值或者属性值的预测可以建模为股票时序图中节点的回归预测问题. 常见目标是预测图中股票节点在下一个时间点或者下一时间段内的价格.

例如, Hou 等人^[65]提出了一个基于股票时序图的深度混合时空图神经网络, 该工作首先通过变分自编码器 (variational auto-encoder, VAE) 对公司之间隐藏关系的降维来构建股票时序图. 然后, 将图卷积网络和 LSTM 集成到一个统一的股价预测网络, 分别抽取用于股票时序图中的时间和空间特征信息, 进行了预测股票价格和股票价格变动趋势两种实验. Pillay 等人^[113]评估和比较了 3 种最先进的时空图神经网络方法, 应用于股价趋势预测领域的单步股价预测和多步股价预测, 发现 GraphWaveNet^[82]在多步预测好于 StemGNN^[84]和 MTGNN^[50], 而 MTGNN 在单步预测具有优势. 本文第 3.1.1 节提到的 Tian 等人^[67]的工作亦可以进行回报率的多步回归预测.

3.1.3 排名列表

现有研究通常将股价趋势预测当作节点分类任务来预测股票涨跌趋势, 或作为一个节点回归任务来预测股票具体价格, 但是无法用于选择预期利润最大的股票, 这可能导致股票投资结果次优. 因此, Feng 等人^[63]综合考虑股票回报率具体数值和股票盈利能力相对排名两种因素, 提出了一种基于时间图卷积 (temporal graph convolution) 网络的股票盈利能力排名预测方法, 名为 RSR (relational stock ranking). 如图 7 所示, 首先将每只股票的历史价格序列数据 \mathcal{X} , 输入到 LSTM 网络, 以捕获序列相关性并学习股价序列的嵌入 E_t . 随后, 将 E_t 送入时间图卷积网络, 该网络可实现等效的图卷积操作, 即使用谱卷积来捕获傅里叶域中的局部连接. 最后, 综合考虑投资回报率和排名预测任务, RSR 使用了一个结合了点回归损失和成对排名感知损失的目标函数, 来优化其模型参数:

$$l(\hat{r}^{t+1}, r^{t+1}) = \|\hat{r}^{t+1}, r^{t+1}\|^2 + \alpha \sum_{i=0}^N \sum_{j=0}^N \max(0, -(\hat{r}_i^{t+1} - \hat{r}_j^{t+1})(r_i^{t+1} - r_j^{t+1})) \quad (2)$$

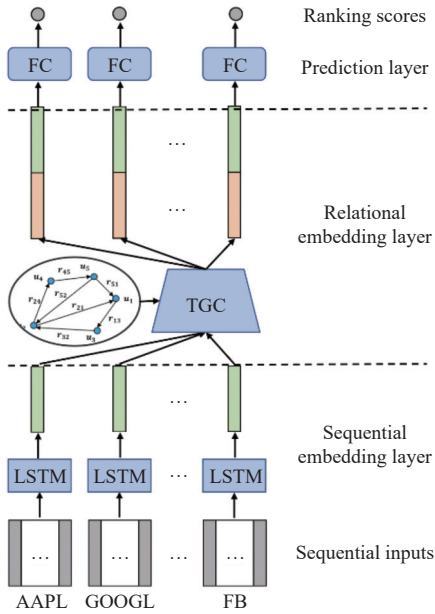


图 7 RSR 的模型图

其中, r^{t+1} 和 \hat{r}^{t+1} 分别为真实值和预测值, α 为平衡两个损失项的超参数. 将提出的综合损失 l 最小化, 可以使排名得分既接近股票收益率绝对数值大小, 又接近股票收益率之间的相对排名顺序.

为了解决不同股票间关系强度随时间发生变化的问题, Ying 等人^[64]对上面提到的关系股票排名算法 RSR 进行了改进, 提出了一种基于时间感知的图关系注意力网络 (time-aware graph relational attention network, TRAN). 首先, TRAN 基于板块-行业分类信息构建股票关系图, 分别从历史股价序列和股票描述文档中提取股价的历史特征和股票属性; 然后, TRAN 通过历史股价特征和股票属性的动态交互来获取股票间的时变关系注意力权重. 最后, 基于所学习的股票关系注意力, TRAN 利用图卷积运算更新相关股票节点的特征表达. 具体地, 股票 s 的股价和其属性在第 t 天的动态交互函数可表示为:

$$e_t^s = \Psi(W_{f_1}^T [\mathbf{h}_t^s; W_{f_2} Z_s] + b_f) \quad (3)$$

其中, Ψ 是激活函数, Z_s 表示从股票描述文档中提取的

股票属性, \mathbf{h}_t^s 表示股票价格历史序列的编码特征, W_{f_2} 、 W_{f_1} 表示全连接层的参数矩阵, T 是转置操作, b_f 表示偏置. 具体地, 股票 s 和 j 在第 t 天的时间感知关系注意力权值的计算可形式化为:

$$\eta_t^{s,j} = \frac{\exp(\alpha_t^s(j))}{\sum_{k \in N_s} \exp(\alpha_t^s(k))} \quad (4)$$

$$\alpha_i^s(j) = u_a^T \phi(W_a e_i^j + b_a) \tag{5}$$

其中, $\alpha_i^s(j)$ 表示股票 j 的特征对股票 s 的重要性, N_s 是股票 s 的邻域, ϕ 为激活函数, u_a^T 、 W_a 和 b_a 是需要学习的参数向量和矩阵. 和一般的 GCN 模型的邻接矩阵是固定的不同, TRAN 可利用关系注意力权重更新关系特征表达. 具体地, 对于股票 s , 将其在图中所有相邻股票节点的特征聚合起来, 得到第 t 天的关系特征表示 \mathbf{a}_i^s . 图的卷积运算公式如下:

$$\mathbf{a}_i^s = \sum_{j \in N_s} \frac{\eta_i^{s,j}}{\text{degree}(j)} \mathbf{h}_i^j \tag{6}$$

其中, $\text{degree}(j)$ 为股票节点 j 的度数, \mathbf{h}_i^j 是最后一个隐藏状态. 最后, 利用图卷积的输出和股票历史特征来预测所有股票的排名.

以上排名预测方法仅关注股票间的成对关系, 忽视了股票之间的高阶关系. 而超边通过连接多个节点将普通图扩展为超图, 由此可以表示节点所属集合间的高阶关系. 为了建模股票的共同变动趋势, 基于超图的排名列表预测逐渐兴起. Sawhney 等人^[93]提出了用于股票排名的时空超图注意网络 STHAN-SR, 其利用 Hawkes 过程^[94]增强了基于 LSTM 提取的时间注意力, 得到时序股票特征. 然后, STHAN-SR 再利用超图卷积将股票超图的空间结构信息嵌入到股票节点的特征表达中. 考虑到股票之间的动态相关性, Sawhney 等^[86]提出时间演进动态股票超图网络 (time evolving dynamic stock hypergraph network, TEDHGN), 以时间敏感的方式, 通过构建基于行业分类、公司关系和股票动态相关性的超图, 学习股票之间的共同股价趋势.

3.1.4 本节总结

表 2 从不同任务角度, 简要总结了第 3.1 节股价趋势风险预测部分提到文章所使用的数据来源、时空信息处理方式和模型名称.

表 2 股价变动风险预测方法总结

任务分类	论文	年份	数据来源	时空信息处理方式	模型名称
节点分类	Deng 等人 ^[81]	2019	股票历史价格序列, 新闻文本和结构化知识数据	基于知识图谱的知识驱动事件嵌入获得空间信息, TCN 处理时序数据	KDTCN
	Wu 等人 ^[66]	2022	沪深300成分股每日股价	用 struc2vec 学习结构特征, 两阶段 RNN 处理时序特征	Price graphs
	Tian 等人 ^[67]	2022	标普500, 沪深300和澳大利亚ASX 300的时序股价	混合注意力编码器学习时间特征和节点特征	HAD-GNN
	Xiang 等人 ^[90]	2022	美国标准普尔500指数和中国沪深300指数的构成实体, 还包括公司关系图	用 Transformer 进行历史价格编码, 然后采用异构图神经网络聚合空间信息	THGNN
	Sawhney 等人 ^[85]	2020	标普500股价数据和 GICS 板块-行业数据	门控时序卷积提取时间特征, 超图卷积提取空间特征	STHGCN
	Cui 等人 ^[75]	2021	A股股票的股价数据, 申万行业分类标准和基金持有关系	用门控循环单元提取时序信息, 输入到超图三重注意力网络提取空间信息	HGTAN
节点回归	Hou 等人 ^[65]	2021	2010年标准普尔100综合指数87家公司的分钟级股票数据	根据VAE计算得到的距离构建空间图结构, 再用GCN-LSTM建模时空特征	ST-Trader
排名列表	Pillay 等人 ^[113]	2021	FTSE/JSE Top 40 指数成分股的日收盘价数据	时空图神经网络模型自适应学习空间结构和时间特征	GraphWaveNet MTGNN StemGNN
	Feng 等人 ^[63]	2019	历史价格数据、板块-行业关系和维基公司间关系	显式领域知识注入, 补充空间信息, 设计时间图卷积模块进行时间感知嵌入传播	RSR
	Ying 等人 ^[64]	2020	股票历史价格序列, 雅虎财经的股票描述文档和板块-行业关系	利用RNN建模股票价格序列信息, 利用GNN方法提取空间信息	TRAN
	Sawhney 等人 ^[93]	2021	纳斯达克, 纽交所和东京交易所股价数据, 维基百科关系数据	采用LSTM和Hawkes过程处理时序信息, 用超图卷积提取空间特征	STHAN-SR
	Sawhney 等人 ^[86]	2022	纽交所, 纳斯达克和A股股价	用注意力门控时序卷积提取时间特征, 用超图卷积提取空间特征	TEDHGN

3.2 贷款违约风险预测

贷款违约是指借贷的公司或个人没有在规定期限内还款而导致的违约行为. 传统基于统计推断和机器学习的贷款违约预测方法假设贷款方的违约行为相互独立, 且其信用特征固定不变. 然而, 中小企业的贷款违约概率不仅取决于其自身的财务状况, 也会受到相关联的企业违约状态的影响. 此外, 企业的信用特征也会随时间发生变化. 因此, 现有研究开始利用时序图神经网络, 从基于借贷事件等各种关系的动态图中学习相关的时序和结构特征. 从图学习的任务角度考虑, 现有的方法通常将贷款违约风险预测问题建模为相关时序图中的节点分类、边分类和链接预测任务, 具体介绍如下.

3.2.1 节点分类

此类方法常基于动态的企业-银行借贷事件构成的图数据, 将企业的贷款违约风险问题建模为图中企业节点的二分类任务. 为了捕捉图中的时序和关系结构特征, 研究人员提出了一系列基于动态借贷图的时序图神经网络. 例如, Cheng 等人^[68]最先提出了一个动态违约预测框架 DDPF, 其在一个端到端的框架中同时利用了贷款行为序列和时序贷款担保图来学习企业的动态贷款特征. 具体地, 如图 8 所示, 输入的时序数据中 Q 代表贷款行为序列, 通常包含一系列贷款特征: 如金额、利息、频率等, G 代表当前时间的担保网络, 表示中小企业在当前时间窗口的网络特征. DDPF 首先将企业在不同时间窗口的 Q 和 G 特征进行连接, 以保留企业当前时间的贷款行为和担保网络的空间结构特征. 随后, 将连接后的混合特征序列送入门控 LSTM, 以学习企业在一段时间内的贷款行为和担保的时序特征. 最后, 各个时间窗口内企业贷款特征的加权和被用于贷款违约的风险预测. Cheng 等人^[69]随后提出的 TDEP 结构与之相似, 在此不再赘述. 此类方法最先考虑到担保网络的图特征可以用于建模贷款违约问题, 然而门控 LSTM 模型在处理图数据方面表现不佳, 故需要运用图神经网络方法针对担保关系进行建模.

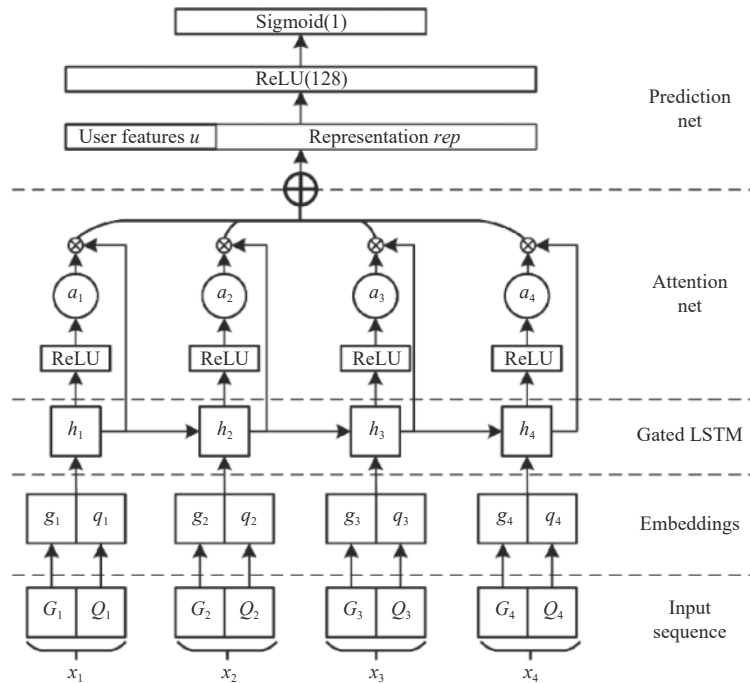


图 8 DDPF 的模型框架图

上述研究仍然面临以下 3 个挑战: 1) 用户贷款时间间隔不规则. 2) 用户的信用风险评级虽然长期内稳定, 但短期内会遭遇突然变化. 3) 图结构和时序信息在同一框架内的整合. 为解决上述问题, Wang 等人^[46]提出了一种基于动态信用图的时间感知图神经网络 (temporal-aware graph neural network, TemGNN) 来预测用户的信用风险, 该网络能够自适应地聚合用户信用风险相关的静态因素、短期因素和长期因素. 具体来说, TemGNN 首先采用多层

感知机 (multilayer perceptron) 学习信用风险相关的静态用户特征. 然后, TemGNN 利用具有特殊图卷积模块的短期图编码器来编码每个图快照, 通过聚合短期的时间和结构特征来丰富用户的特征表达. 与以往 GCN 模型只聚合当前时间的邻居信息不同, TemGNN 假设节点的当前表示也受到最近时间步内其邻居信息的很大影响, 则用户节点 u 的隐式特征表达的学习可形式化为:

$$\hat{\mathbf{h}}_{\mathcal{N}_t(u)} = \sum_{t' \in \text{Prev}_K(t)} \sum_{v \in \mathcal{N}_{t'}(u)} \alpha(u, v, t') \hat{\mathbf{h}}_{v, t'} \quad (7)$$

其中, $\text{Prev}_K(t)$ 代表时刻 t 的图快照和之前的 K 个图快照的集合, $\mathcal{N}_t(u)$ 是节点 u 在时刻 t 的邻居节点, $\alpha(u, v, t')$ 是不同时刻的节点间注意力系数. $\hat{\mathbf{h}}_{v, t'}$ 是节点 v 在时刻 t' 的输入特征. 最后, 连接用户静态财务特征 $\hat{\mathbf{h}}_u$ 和邻居节点特征 $\hat{\mathbf{h}}_{\mathcal{N}_t(u)}$ 来更新节点 u 下一层的用户特征:

$$\hat{\mathbf{h}}_u^{(1)} = \text{ReLU}\left(W_{\text{GNN}}^{(1)} \left(\hat{\mathbf{h}}_u \parallel \hat{\mathbf{h}}_{\mathcal{N}_t(u)}\right)\right) \quad (8)$$

其中, \parallel 代表向量拼接操作, W_{GNN} 表示可训练的权重矩阵, 将多个这样的图卷积模块进行叠加, 形成层次图表示学习模型. 第 3 步, 采用基于间隔衰减注意力的 LSTM 自适应提取动态图中的长期依赖信息.

虽然现有方法能对时序图在每个时间步的嵌入进行加权聚合, 但现有的时间注意力机制背后仍然假设时间步的间隔是均匀分布的. 然而, 实际上两个连续借贷事件的间隔通常从天到年不等. 直观上看, 间隔较大的事件对后续事件的影响较小. 为了考虑这种区间的不规则性, TemGNN 进一步定义了如下的间隔单调递减函数:

$$g(\delta_t) = \frac{1}{\log(e + \delta_t)} \quad (9)$$

其中, δ_t 为时刻 t 到现在的时间间隔. 基于间隔衰减的时间注意力权重, 用户的特征表达学习可被重新定义为:

$$\mathbf{h} = \sum_{i=0}^T (g(\delta_{t_i}) \cdot \alpha'(t_i) \cdot \mathbf{h}_{t_i}) \quad (10)$$

其中, $g(\delta_{t_i})$ 是衰减权重, $\alpha'(t_i)$ 是归一化注意力相关系数. \mathbf{h} 不仅表示当前时间步的用户状态, 而且收集了之前的一些记忆, 并通过间隔衰减注意力机制将这些记忆规则化. 最后, 计算风险评分并执行节点分类预测任务.

3.2.2 边分类

这类方法认为高风险的担保行为是贷款违约事件的导火索, 贷款违约事件会沿错综复杂的担保网络像病毒般蔓延扩散, 故预测担保网络中的贷款违约传染链路很有必要. 因此, 这类方法从传染链 (边) 是否具有风险的角度, 将贷款违约风险的预测问题转化成时序图中边的分类问题. 针对传染链风险评分研究的空白, Cheng 等人^[76]率先提出了一种名为时间注意力传染链增强评分模型 TRACER 的传染链风险评级方法, 其利用时间注意力机制和递归评级神经网络, 从连续贷款行为中推断传染链的风险概率. TRACER 首先根据贷款记录和担保网络生成传染链, 并采用注意力机制学习传染链间的相对重要性. 通过链间注意力层, TRACER 可有效编码空间信息, 解决了传染链数据不足情形下的风险评分问题. 然后, 通过引入 GRU 和时序注意力层来提取时序信息, TRACER 从连续的贷款行为中推断出传染链的风险概率.

虽然取得了不错的预测结果, 但上述工作忽视了对系统性金融风险的研究. 为此, Cheng 等人^[114]重点关注由担保网络中贷款违约衍生出的系统性金融风险的量化研究, 提出了第 1 个针对贷款网络进行系统性金融风险管理的工具 iConReg. 如图 9 所示, iConReg 首先利用贷款网络上的广度优先搜索算法来构建所有可能的传染链. 然后, 由于传染链的风险是由节点信息和传染链的状态决定的, 故 iConReg 利用链内和链间注意力图神经网络将传染链和贷款网络的潜在风险信息嵌入节点的特征表达. 最后, 将所有贷款网络、传染链和节点的特征表达连接后送入卷积和池化网络层预测系统性的金融风险.

针对贷款网络中担保关系会随时间动态变化的问题, 图神经网络需要直接从网络中提取暂时性存在的隐藏模式信息, 并随着网络的更新动态调整其重要性. 因此, Cheng 等人^[77]提出了一种基于时序图的注意力神经网络 DGANN, 其将每个风险担保建模为动态贷款网络中的一条边, 从而将贷款违约的风险建模为边的分类问题. DGANN 使用了一个时序自注意力层来进一步捕捉动态网络中的时间模式信息. 该自注意力层的输出是边在时刻 t 的特征表达 \mathbf{u}_t , 其计算可形式化为:

$$\mathbf{u}_t = \left(\parallel_{j=1}^T \beta_{t,j} W_j \mathbf{h}_j\right) \quad (11)$$

其中, \mathbf{h}_j 是 GRU 在时刻 j 的输出, $\beta_{t,j}$ 是时刻 t 和 j 之间的注意力系数, W_t 是训练的参数矩阵, \mathbf{u}_t 是最终的输出, 符号 \parallel 代表向量的拼接操作. 最后, 将 \mathbf{u}_t 输入到预测层计算担保边的风险概率.

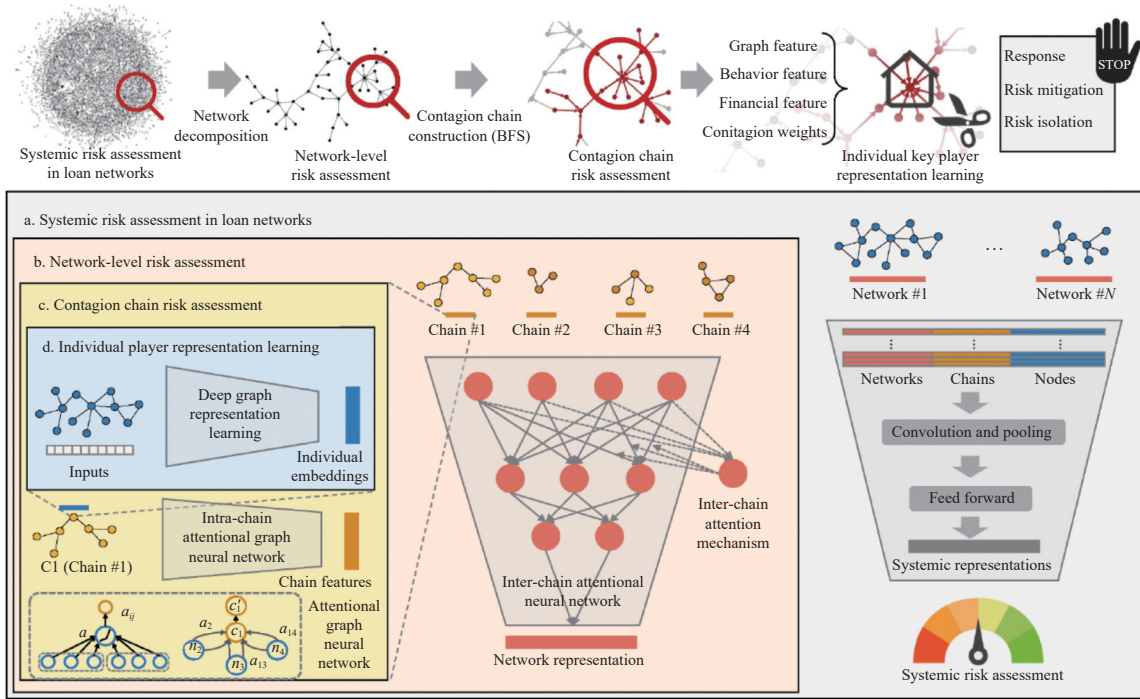


图 9 iConReg 的模型框架图

3.2.3 链接预测

针对中小企业信用风险预测相关数据缺乏的问题, 这类方法将企业贷款风险的预测问题建模为时序贷款网络中企业的节点分类和链接预测的联合学习任务. 例如, Yang 等人^[70]提出了一种时空感知的图神经网络, 其利用基于时序贷款图中企业节点间的供应关系预测任务, 来提升企业节点贷款风险分类预测结果. 具体而言, 如图 10 所示, Yang 等人通过所提出的时空感知图神经网络 ST-GNN, 将图中企业间的供应链预测任务和企业节点的贷款风险分类任务集成到了一个端到端的网络框架中. 首先, ST-GNN 基于已有的企业贷款信用数据, 进行不同时刻企业间的供应链挖掘, 并在此基础上, 使用一个基于注意力的空间感知聚合模块, 学习每个企业节点在不同时刻贷款图快照的空间结构特征. 然后, ST-GNN 使用一个类似 LSTM 的时间感知聚合模块建模图快照序列的时序变化信息. 最后, 通过一个注意力机制, ST-GNN 将所学的空间和时间特征相结合, 用于企业的贷款风险评估. 综上, ST-GNN 能够自适应地从 T 个时刻的图快照中节点 u 的邻域捕获空间和时间信息, 用于最后的贷款风险预测.

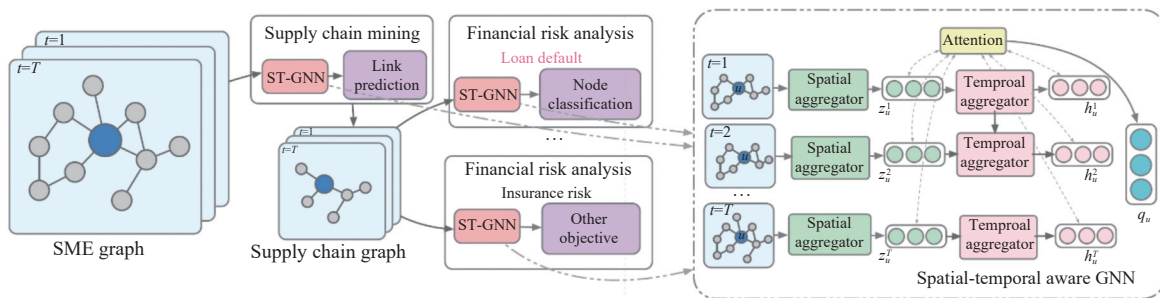


图 10 ST-GNN 的模型框架图

3.2.4 本节总结

表3从不同任务角度,简要总结了贷款违约风险预测部分提到文章所使用的数据来源、时序信息处理方式和模型名称,该部分任务可以被建模为节点分类、边分类和链接预测这3种。

表3 贷款违约风险预测方法总结

任务分类	论文	年份	数据来源	时序信息处理方式	模型名称
节点分类	Cheng等人 ^[68]	2019	中小企业4年内的贷款记录	门控LSTM和注意力机制捕获时间信息	DDPF
	Cheng等人 ^[69]	2020	东亚一家主要的商业银行一年内的贷款记录	门控LSTM和注意力机制捕获时间信息	TEDP
	Wang等人 ^[46]	2021	支付宝用户的借贷记录	LSTM捕获时间信息	Tem-GNN
边分类	Cheng等人 ^[76]	2020	中国一家主要商业银行2013–2016年的贷款数据	GRU时序注意力层	TRACER
	Cheng等人 ^[114]	2020	从长三角和珠三角的商业银行获得一个全面的贷款数据集	暂未使用时序信息	iConReg
	Cheng等人 ^[77]	2020	2013–2016东亚主要银行的贷款记录	GRU结合自注意力机制处理时序信息	DGANN
链接预测	Yang等人 ^[70]	2021	蚂蚁中小企业贷款数据和供应链关系	类似LSTM时间感知聚合器处理时序信息	ST-GNN

3.3 欺诈交易风险预测

欺诈交易指的是未经授权、隐瞒身份和交易意图的交易,使受害者遭受大量财产损失。最早的欺诈交易集中于银行卡盗刷等基于身份盗窃的行为。随着互联网的发展,欺诈交易产生线上化、平台化的发展趋势。传统的基于规则和机器学习的欺诈交易预测方法,直接将人工生成的潜在风险特征输入到分类模型中,而忽略了对欺诈行为空间和时间模式特征的联合学习。此外,这些传统方法不适用于现实世界中具有复杂和不可预测的欺诈模式的大规模欺诈检测事件。通过研究发现,欺诈交易行为具有时间特征,如果在用户最近的交易行为中观察到异常模式,则欺诈交易的可能性很高;欺诈交易行为同时具有空间特征,体现出地理聚集性。由于交易生态系统构成一个由用户与商家交互的巨大网络,因此可以捕捉该网络结构的各种信号。因此,近年来有不少学者基于时序图神经网络展开欺诈交易的预测研究。

3.3.1 节点分类

这类方法通常将交易记录视为边,将参与交易的实体看作节点,通过交易实体节点的二分类任务判断当前节点是否是欺诈交易者。例如,Rao等人^[88]从可疑的大规模账号注册行为入手分析,认为这种滥用行为容易滋生后续欺诈交易,他们设想通过大规模可疑注册来预测未来的欺诈交易,于是提出动态异构图神经网络DHGReg。如图11所示,DHGReg通过注册账号的结构子图和时序子图上的结构和时间卷积,来捕获滥用账户关联的空间特征和最近注册账户的时间特征。

然而,这种方法从滥用注册行为入手间接分析,并没有直接针对欺诈交易的行为进行建模。针对此问题,在分析可疑大规模注册基础上,Rao等人^[71]直接针对标记风险交易和识别风险账户的应用场景,提出了动态异构图神经网络DyHGN及其变体网络,其通过不同的GCN层实现结构信息和时间信息建模。

为了更好地捕获欺诈交易行为的时间和空间特征,Reddy等人^[96]提出了一个基于时间点过程和图神经网络的欺诈检测框架TeGraF,其联合利用基于交易序列的用户时间特征和基于交易网络的结构特征来训练欺诈交易分类器。具体而言,TeGraF一方面采用时间点过程(temporal point processes)捕获用户交易的时间信息;另一方面,使用图神经网络模型对用户和商家之间的关系进行了建模,从而学习了交易的空间结构信息。最后,TeGraF将所学节点的时间和图结构特征连接后,送入风险分类模块以进行欺诈交易预测。

当前,时序图神经网络在建模风险的时序特征时仍然存在如下两个问题:1)提出时序特征时可能用到未来时间的特征信息。2)两跳以上图结构的推理效率低下。为了解决上述问题,Lu等人^[95]提出了面向欺诈检测批和实时图拓扑框架BRIGHT,其由一个两阶段有向图,以及支持批处理和实时处理的Lambda神经网络(LambdaNetworks)^[97]构成。两阶段有向图由分别代表历史关系和实时链接的两个子图构成,它能保证邻居节点传递的信息来自历史的

交易数据. 所用的 LambdaNetworks 将推理分解成了两个步骤: 批和实时的交易预测推理. 通过快照聚合技术, LambdaNetworks 能够避免在网络训练期间使用到节点在未来的特征信息. Lambda 架构将训练过程分为离线训练和在线推理两个模块, 实现了高效的在线实时推理并部署到工业系统中, 解决了 DHGReg 和 DyHGN 的内存消耗大和查询延迟高的问题. 针对图神经网络推理效率低的问题, Wang 等人^[92]提出了异步传播注意力网络 APAN, 该模型由同步推理和异步传播模块构成. 其中, 前者采用基于多头注意力机制的编码器和多层感知机解码器架构. 在后者中, 一种新的邮箱机制被用来收集并保存 k-hop 邻居节点沿时间戳边传播的信息. 通过图推理与图计算的解耦, APAN 能有效减轻繁重的图查询操作对模型推理速度的损害, 极大地加速了在线的实时图推理.

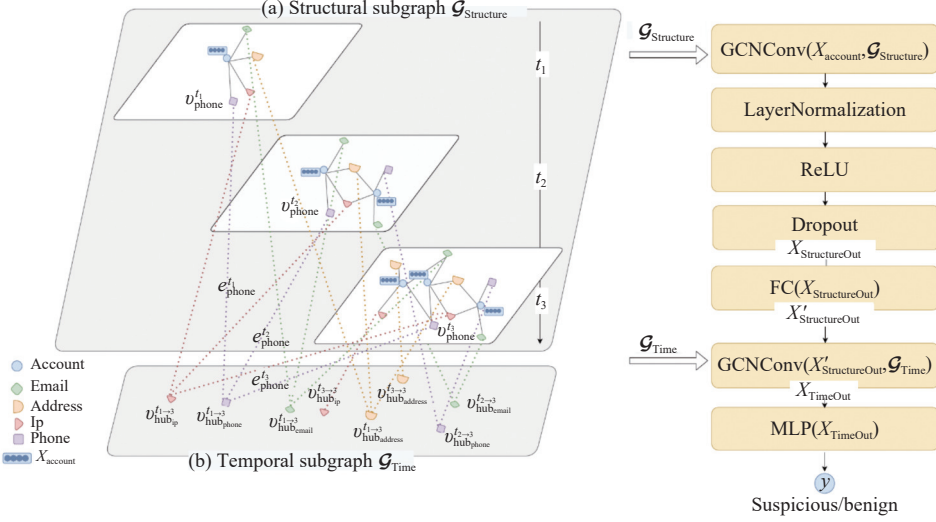


图 11 DHGReg 的模型框架图

针对现有金融异常检测数据集忽略背景节点、规模太小和不含有时序信息的问题, Huang 等人^[115]提出了大型借贷网络数据集 DGraph, 其是一个由超过 300 万个节点和 400 万条边组成的大规模动态图. 该数据集提供了超过 100 万个标注节点, 此外, 保留了 200 万以上的无标注背景用户节点, 它们代表缺乏借贷行为的非检测目标用户. 在该数据集中, 异常用户和正常用户在网络结构、邻居特征分布和时间动态方面具有不同的特征. 目前流行的监督学习方法和无监督学习方法在该数据集上的性能还有很大提升空间.

3.3.2 边分类

在金融实体交易的图数据中, 边表示两个金融实体间的交易行为. 此类方法的目标是根据给定时间内的交易记录, 推断新产生的交易是否是欺诈事件. 它们把某交易是否是欺诈交易的预测任务建模为动态交易图中边的分类任务. 例如, Cheng 等人^[89]提出一种新的信用卡欺诈交易检测模型 STAGN. 如图 12 所示, STAGN 使用时空注意力机制有效提取时间和空间信息整合成三维张量, 然后通过引入卷积操作来捕获时空模式之间的相互依赖关系. 具体地, STAGN 首先利用基于位置的图神经网络层提取信用卡交易图的结构特征. 然后, STAGN 再按顺序提取时间和空间注意力, 其计算过程可形式化为:

$$rept = \sum_{\bar{i}=1}^{N_1} a_{1,\bar{i}} \mathcal{X}(\bar{i}, :, :)$$
 (12)

$$\mathcal{H}^a = \sum_{\bar{s}=1}^{N_2} a_{2,\bar{s}} rept(:, \bar{s}, :)$$
 (13)

$$a_{1,\bar{i}} = \frac{\exp((1 - \lambda_1) \cdot NN_{\bar{i}}(W_{\bar{i}}, \mathcal{X}(\bar{i}, :, :)))}{\sum_{\bar{i}=1}^{N_1} \exp((1 - \lambda_1) \cdot NN_{\bar{i}}(W_{\bar{i}}, \mathcal{X}(\bar{i}, :, :)))}$$
 (14)

$$a_{2,\bar{s}} = \frac{\exp((1-\lambda_2) \cdot NN_{\bar{s}}(W_{\bar{s}}, rept(\cdot, \bar{s}, \cdot)))}{\sum_{\bar{s}=1}^{N_2} \exp((1-\lambda_2) \cdot NN_{\bar{s}}(W_{\bar{s}}, rept(\cdot, \bar{s}, \cdot)))} \quad (15)$$

其中, $rept$ 是时序注意力层的输出, \mathcal{H}^a 表示空间注意力层的输出, $a_{1,t}$ 和 $a_{2,\bar{s}}$ 是每个时间片和空间片的权重, $\mathcal{X}(\bar{t}, \cdot, \cdot)$ 和 $rept(\cdot, \bar{s}, \cdot)$ 是交易记录张量 \mathcal{X} 的时间和空间特征. 随后, 利用三维卷积处理时空信息, 可表示为:

$$\mathcal{H}_j^c = \sum_j \mathcal{H}^{c-1}(c_t - c_m, c_s - c_n, c_f - c_o) \mathcal{W}_i^c(c_m, c_n, c_o) \quad (16)$$

其中, \mathcal{W}_i^c 是 3D 卷积核 (c_m, c_n, c_o) 的权重. 最后, STAGN 将得到的特征向量输出到检测层进行评分预测, 以显示当前交易是欺诈交易的概率.

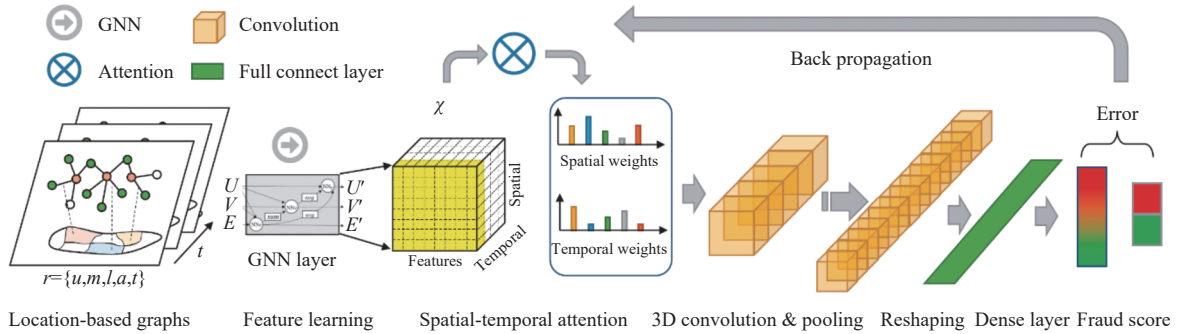


图 12 STAGN 的模型框架图

3.3.3 本节总结

表 4 按照任务类型, 简要总结了欺诈交易风险预测部分提到文章所使用的数据来源、时序信息处理方式和模型名称, 欺诈交易风险预测部分的时序图学习任务可以被建模为节点分类和边分类两种.

表 4 欺诈交易风险预测方法总结

任务类型	论文	年份	数据来源	时序信息处理方式	模型名称
节点分类	Rao等人 ^[88]	2020	电子商务平台的账户注册记录	根据不同时间的静态结构快照构建时间子图	DHGReg
	Rao等人 ^[71]	2022	eBay的账户注册记录和交易记录	根据不同时间的静态结构快照构建时间子图	DyHGN
	Reddy等人 ^[96]	2021	由真实交易数据合成生成	用不同TPP算法捕获时间过程	TeGraF
	Lu等人 ^[95]	2022	eBay的真实交易记录	LambdaNetworks进行批处理和流处理, 后者用于处理实时图	BRIGHT
	Wang等人 ^[92]	2021	AliPay数据	不同时刻到达邮箱的邮件进行位置编码	APAN
边分类	Huang等人 ^[115]	2019	Finvolution平台的借款记录	构建了金融领域大型动态图	DGraph
	Cheng等人 ^[89]	2020	商业银行信用卡交易记录	采用时间注意力层捕获时间片特征	STAGN

3.4 洗钱和逃税风险预测

洗钱是犯罪分子试图掩盖其非法所得财务的性质和来源的一种行为, 不仅会影响金融市场的稳定, 严重影响经济的健康发展, 甚至会助长更大规模的犯罪活动, 严重危害人民生命财产安全和国家的长治久安. 因此, 开展“反洗钱”活动, 科学合理地预测洗钱风险十分必要. 近年来, 随着互联网的发展, 天然具有匿名性、去中心化特点的比特币^[116]等加密货币成为洗钱行为的新载体. 基于比特币等加密货币的洗钱行为, 给“反洗钱”追踪带来困难, 使得传统的基于统计学和机器学习的方法捉襟见肘. 同样, 逃税行为的危害也应引起重视. 现有研究通过将现金流关系构建为图, 运用图神经网络方法学习图中账户节点间的时间关系和空间拓扑结构, 来捕捉洗钱和逃税行为特征.

3.4.1 基于节点分类的洗钱风险预测

由于洗钱活动往往涉及实体之间的现金流关系, 基于图的分析已成为“反洗钱”策略中日益重要的工具. 例如, 可以构造以单个账户为节点, 两个账户之间的现金流关系为边的图. 现有研究一般将洗钱风险的预测问题建模为现金流图中节点的分类问题, 目的是预测图中账户节点涉嫌洗钱活动的可能性. 例如, Weber 等人^[117]通过一个合成的动态现金流图数据测试了现有的图神经网络模型 (例如, GCN 和 FastGCN^[118]) 在“反洗钱”任务上的性能. 在该现金流图中, 每个节点表示一个账户, 且节点有包括账号, 账户类别和创建日期等属性信息. 账号之间的交易被表为带有时间戳的边, 这些边被稀疏地标记为可疑交易和确认可疑交易. 该任务预测给定目标账户节点的可疑性, 并通过与已知可疑节点的直接或间接连接识别交易网络中其他潜在的不良账户.

针对现有“反洗钱”研究缺乏大规模真实可用数据集的问题, Weber 等人^[119]随后贡献了一个称为 Elliptic 的比特币交易时间序列图数据集, 最初包含 203 769 个节点交易, 234 355 个定向边支付流和 166 种丰富的包括基于非公开数据提取的节点特征.

针对传统 GCN 无法有效提取洗钱行为的时间和空间特征的问题, Xia 等人^[72]提出了一种基于图卷积神经网络和 LSTM 混合的时空洗钱预测模型 MGC-LSTM, 其首先利用 LSTM 提取洗钱交易数据的时间依赖特征, 然后利用 GCN 学习不同洗钱交易之间的空间依赖性. 具体地, 如图 13 所示, 给定当前时间戳的交易序列和图数据, 该模型首先使用 GCN 学习该交易图数据的空间结构特征, 然后将所学特征送入 LSTM 来处理不同图快照的空间特征序列. 之后, 将当前时间戳网络的训练结果输出到下一个时间戳的网络循环训练, 以此来捕捉洗钱活动的时空依赖.

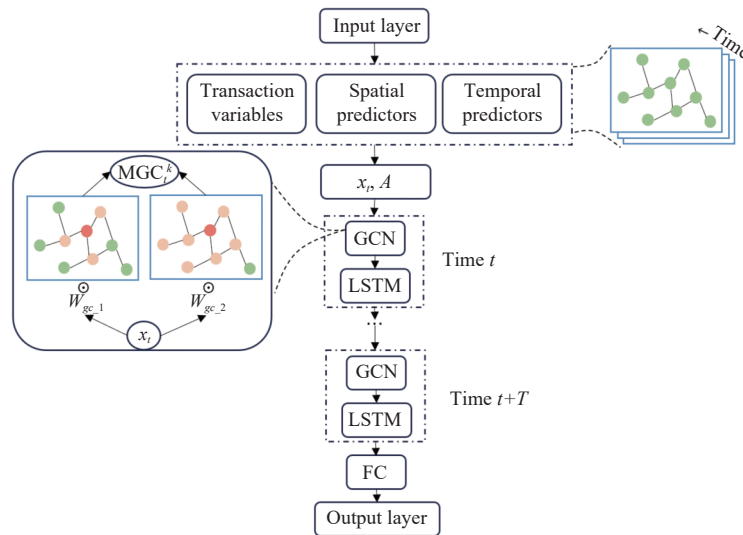


图 13 MGC-LSTM 的模型框架图

针对现有的反洗钱研究极少考虑交易时间的问题, Alarab 等人^[73]开发了一个将 LSTM 与 GCN 相结合的分类模型. 在该模型中, LSTM 被用来建模比特币交易图的时间序列, GCN 聚合目标节点的前 k 个有影响力的邻居节点信息, 并据此建模结构特征. 此外, Alarab 等人引入了一个应用于大规模比特币交易图数据集的主动学习 (active learning) 框架解决了数据匮乏和洗钱交易数据标注成本高的问题. 该框架的标注器使用采集函数从未标记的训练集中查询数据标签, 自动学习查询信息量最大的数据的标签, 空白数据标注完成后放入训练集并再次训练模型.

针对现有的基于规则的“反洗钱”系统存在风险排序、假阴性和未考虑交易网络结构信息的 3 个主要限制, Assumpcao 等人^[99]提出了一个多任务学习框架 DELATOR 来检测动态金融交易网络中的洗钱行为. 第 1 步, 在每个由交易数据和原始特征构建的离散时间快照图上引入多任务学习, 用传统图神经网络建模空间信息, 将不同快照获得的嵌入按时间顺序拼接以包含时间信息. 具体而言, 对于每个图快照, 基于非监督链接预测任务和监督边回归任务生成丰富的节点嵌入, 以此来充分利用网络中不同的信息. 第 2 步, 采用 GraphSMOTE^[120]来解决现有数据

集不平衡的问题, 最后训练一个监督分类器来预测目标客户参与洗钱的概率。

3.4.2 基于节点分类的逃税风险预测

逃税是指纳税人通过欺骗、隐瞒等手段逃避纳税的违法行为。逃税风险预测一般建模为节点分类问题, 预测目标节点是否参与从事偷税漏税活动。传统的逃税风险预测方法采用基于人工生成的特征训练分类器, 但它们具有严重依赖专家经验和不适用于动态交易网络的缺点。为了克服这些挑战, Wang 等人^[79]提出了一种新的时序边增强的图注意力网络逃税检测方法。如图 14 所示, 该方法将边增强图注意力网络 (edge enhanced graph attention network, EGAT) 和循环加权平均单位 (recurrent weighted average unit, RWA) 相结合。具体来说, 利用 EGAT 捕获 t 时刻交易网络的复杂拓扑结构, 获取空间特征; 其次, 将获得的具有空间特征的时间序列输入 RWA, 通过单元间的信息传输获取时间特征; 最后, 将所获得的特征向量输入到分类器以预测逃税风险。该模型不仅利用了节点的基本特征, 还考虑了交易网络的动态变化和多维的边特征, 更符合逃税检测的场景。

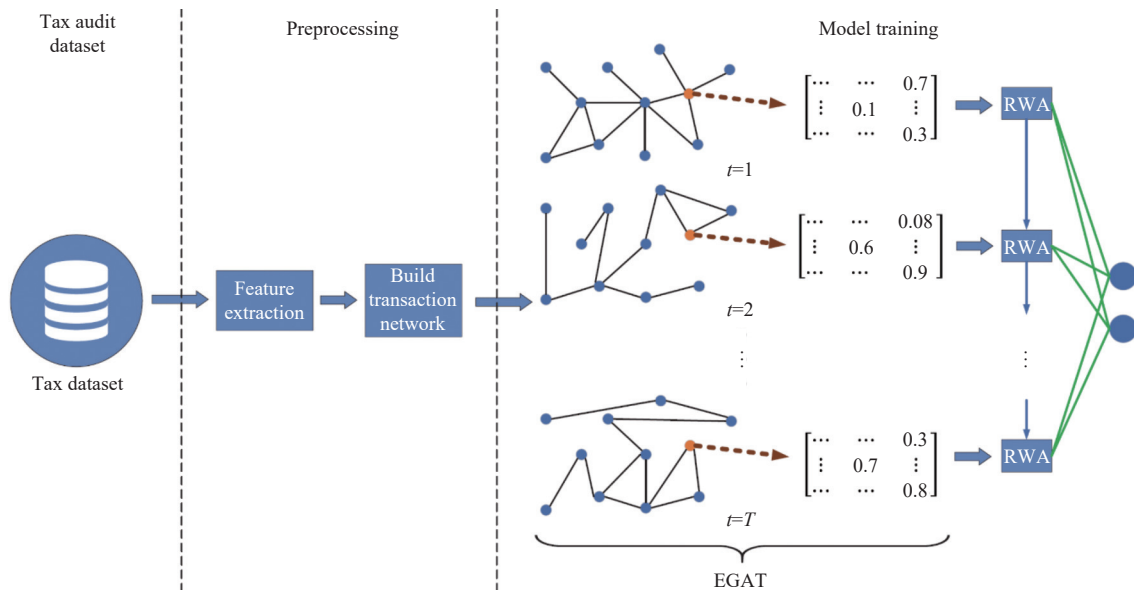


图 14 逃税检测模型流程图

3.4.3 本节总结

表 5 总结了洗钱和逃税风险预测部分提到本文所使用的数据来源、时序信息处理方式和模型名称, 该部分任务总体上均被建模为节点分类。

表 5 面向洗钱和逃税风险预测的节点分类任务总结

论文	年份	数据来源	时序信息处理方式	模型名称
Weber等人 ^[117]	2018	模拟数据	暂未考虑时间信息	AMLSim
Weber等人 ^[119]	2019	Elliptic公司提供的比特币交易数据	使用EvolveGCN建模时间信息	Elliptic数据集
Xia等人 ^[72]	2022	Elliptic数据集	不同时间快照图用LSTM处理时间信息	MGC-LSTM
Alarab等人 ^[73]	2022	Elliptic数据集	不同时间快照图用LSTM处理时间信息	Temporal-GCN
Assumpcao等人 ^[99]	2022	Elliptic数据集	对每个快照的嵌入使用简单的拼接以建模时间信息	DELATOR
Wang等人 ^[79]	2020	中国某省2016–2018年的税务数据	不同时间快照图用RWA处理时间信息	T-EGAT

4 挑战与展望

本文回顾了时序图神经网络在金融风险预测领域的应用。应当注意到, 时空图神经网络最早用于解决交通流

量预测和人体动作识别问题,后来逐渐发展到金融风险预测领域.随着研究的不断深入,现已在离散金融时序图方向有较多应用,在连续金融时序图方面正逐步探索.时序图神经网络和金融风险预测任务的结合是一个令人鼓舞的新方向,但也存在如下挑战和困难.

(1) 缺乏公开的大规模金融风险预测数据集

金融数据高度敏感,与个人隐私、国家机密密切相关,真实的个人借贷、欺诈交易数据集不容易获得,即使相关产业内部研究人员能够获取到的交易数据也需要做脱敏处理,其数据集由于隐私保护等原因难以公开,导致大型的金融图数据难以获得,这阻碍了金融风险预测领域的时序图神经网络的研究.目前,已有部分研究从合作公司处取得匿名或脱敏的数据,完成大规模欺诈交易风险预测金融数据集的构建^[115,119,121],其他金融风险场景的数据集仍待研究者们持续贡献.

(2) 捕获时间、空间特征的建模方法还需精心设计

真实的金融网络较为复杂,可能具有动态异构的节点和边,以及多维度的节点和边属性.同时,这些节点间可能同时存在多种复杂的高阶关联.现有大多数模型依赖于图快照,先在每个快照用图神经网络方法提取时空结构信息,然后用时序方法捕获图快照之间的时序信息.然而,这种基于图快照的方法只能捕捉粗粒度的时序特征.因此,面对动态变化且结构复杂的金融时序图数据,如何对其中蕴含的连续演化过程进行建模还是一个挑战.如何将其他领域的方法^[37-42]迁移应用到金融领域,进行连续金融时序图建模的研究,是潜在的研究热点之一.

(3) 小样本环境下的金融风险预测研究

近年来,小样本学习从图像分类领域兴起,现已应用到时序图神经网络领域.小样本学习是指在数据很少或精心标注的数据很少的情况下进行学习.金融数据之所以存在小样本情形,是因为以下两方面原因,一是出于保护用户金融数据隐私和安全角度考虑,获取到真实、大量的金融数据是困难的,二是动用人工对大规模的数据集精心标注耗时费力,成本极高.虽然图神经网络拥有半监督学习的优势,可以通过边来聚合不可见的节点信息从而完成归纳学习,能在一定程度上缓解小样本数据标注不足的情况,但若若要最大限度发挥模型在小样本环境下的性能、提高建模精确度和通过知识迁移解决“冷启动”等问题,还需要与元学习、知识蒸馏、自监督、主动学习、对比学习等最新方法结合.

(4) 训练和预测的实时性有待提高

金融风险复杂多变,预测、防范金融风险的时机转瞬即逝,需要发展在线实时推理技术,在用户可以忍受的响应延迟内给出风险判断和应对策略.真实金融网络的高频更新对图神经网络模型的响应时间提出了严格要求.然而,图神经网络在聚合目标节点周围多跳的领域信息时,时延明显增加.如何结合在线学习、在线推理技术,在极短时间内得到新的节点或者边的嵌入特征,是未来重要的研究方向之一.

(5) 模型的可解释性有待研究

随着图神经网络在金融风控等领域的应用,对图神经网络可解释性的要求也越来越高^[122].目前可解释性研究在视觉和文本领域已取得重大进展,一些研究者从模型或梯度等角度提供了图神经网络可解释性的研究思路.然而,面向金融风险预测的时序图神经网络的可解释性研究目前还很不完善,这是一个极具潜力的研究方向.

5 全文总结

近年来,基于时序图神经网络的金融风险预测已成为研究热点,近年来已涌现出大量相关工作.本文是第1个对时序图神经网络在金融风险预测领域的研究进行综述的文章,总结和归纳了相关时序图神经网络的特点,以及各类方法的数据来源和时序数据的建模方法,相信可以为未来的研究工作提供宝贵的基础知识和经验.具体地,本文首先介绍了离散和连续时序图的区别,并且根据对时间信息的不同建模方式,将近年来相关的时序图神经网络研究分为基于RNN,基于CNN和基于注意力机制这3大类.可以看到,现有方法多是将图神经网络和RNN结合,用图神经网络提取空间结构信息,再输入到RNN的每个循环单元学习时间信息;或是首先对输入序列用RNN方法捕获时间信息,然后与时序图的节点间关系信息结合,最后再执行图上的空间卷积.随后,为了更清晰地认识时

序图神经网络执行金融风险预测任务的特点, 本文围绕股价趋势风险、贷款违约风险和欺诈交易风险等 4 大类金融风险, 对现有的时序图神经网络按照图学习任务类别分别进行了总结和阐述. 例如, 各类金融风险的预测问题可建模为基于金融时序图的节点分类、节点回归、边分类和链接预测等图学习任务. 最后, 本文指出了当前基于时序图神经网络的金融风险预测的挑战, 并对未来潜在的研究方向进行了展望.

References:

- [1] Kuester K, Mittnik S, Paolella MS. Value-at-risk prediction: A comparison of alternative strategies. *Journal of Financial Econometrics*, 2006, 4(1): 53–89. [doi: [10.1093/jjfnec/nbj002](https://doi.org/10.1093/jjfnec/nbj002)]
- [2] Peng Y, Wang GX, Kou G, Shi Y. An empirical study of classification algorithm evaluation for financial risk prediction. *Applied Soft Computing*, 2011, 11(2): 2906–2915. [doi: [10.1016/j.asoc.2010.11.028](https://doi.org/10.1016/j.asoc.2010.11.028)]
- [3] Valaskova K, Kliestik T, Svabova L, Adamko P. Financial risk measurement and prediction modelling for sustainable development of business entities using regression analysis. *Sustainability*, 2018, 10(7): 2144. [doi: [10.3390/su10072144](https://doi.org/10.3390/su10072144)]
- [4] Gama APM, Geraldies HSA. Credit risk assessment and the impact of the new basel capital accord on small and medium-sized enterprises: An empirical analysis. *Management Research Review*, 2012, 35(8): 727–749. [doi: [10.1108/01409171211247712](https://doi.org/10.1108/01409171211247712)]
- [5] Huang YP, Zhang LM, Li ZH, Qiu H, Sun T, Wang X. Fintech credit risk assessment for SMEs: Evidence from China. 2020. <https://ideas.repec.org/p/imf/imfwpa/2020-193.html>
- [6] Kou G, Xu Y, Peng Y, Shen F, Chen Y, Chang K, Kou SM. Bankruptcy prediction for SMEs using transactional data and two-stage multiobjective feature selection. *Decision Support Systems*, 2021, 140: 113429. [doi: [10.1016/j.dss.2020.113429](https://doi.org/10.1016/j.dss.2020.113429)]
- [7] Gong YC, Jia YQ, Leung T, Toshev A, Ioffe S. Deep convolutional ranking for multilabel image annotation. In: *Proc. of the 2nd Int'l Conf. on Learning Representations. Banff: ICLR*, 2014.
- [8] Oquab M, Bottou L, Laptev I, Sivic J. Learning and transferring mid-level image representations using convolutional neural networks. In: *Proc. of the 2014 IEEE Conf. on Computer Vision and Pattern Recognition. Columbus: IEEE*, 2014. 1717–1724. [doi: [10.1109/CVPR.2014.222](https://doi.org/10.1109/CVPR.2014.222)]
- [9] Song LY, Liu J, Qian BY, Sun MX, Yang K, Sun M, Abbas S. A deep multi-modal CNN for multi-instance multi-label image classification. *IEEE Trans. on Image Processing*, 2018, 27(12): 6025–6038. [doi: [10.1109/TIP.2018.2864920](https://doi.org/10.1109/TIP.2018.2864920)]
- [10] Durand T, Mordan T, Thome N, Cord M. WILDCAT: Weakly supervised learning of deep ConvNets for image classification, pointwise localization and segmentation. In: *Proc. of the 2017 IEEE Conf. on Computer Vision and Pattern Recognition. Honolulu: IEEE*, 2017. 5957–5966. [doi: [10.1109/CVPR.2017.631](https://doi.org/10.1109/CVPR.2017.631)]
- [11] Tang P, Wang XG, Bai S, Shen W, Bai X, Liu WY, Yuille A. PCL: Proposal cluster learning for weakly supervised object detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2020, 42(1): 176–191. [doi: [10.1109/TPAMI.2018.2876304](https://doi.org/10.1109/TPAMI.2018.2876304)]
- [12] Song LY, Liu J, Sun MX, Shang XQ. Weakly supervised group mask network for object detection. *Int'l Journal of Computer Vision*, 2021, 129(3): 681–702. [doi: [10.1007/s11263-020-01397-w](https://doi.org/10.1007/s11263-020-01397-w)]
- [13] Ren ZZ, Yu ZD, Yang XD, Liu MY, Lee YJ, Schwing AG, Kautz J. Instance-aware, context-focused, and memory-efficient weakly supervised object detection. In: *Proc. of the 2020 IEEE/CVF Conf. on Computer Vision and Pattern Recognition. Seattle: IEEE*, 2020. 10595–10604. [doi: [10.1109/CVPR42600.2020.01061](https://doi.org/10.1109/CVPR42600.2020.01061)]
- [14] Goyal Y, Khot T, Summers-Stay D, Batra D, Parikh D. Making the V in VQA matter: Elevating the role of image understanding in visual question answering. In: *Proc. of the 2017 IEEE Conf. on Computer Vision and Pattern Recognition. Honolulu: IEEE*, 2017. 6325–6334. [doi: [10.1109/CVPR.2017.670](https://doi.org/10.1109/CVPR.2017.670)]
- [15] Wang P, Wu Q, Shen CH, Dick A, van den Hengel A. FVQA: Fact-based visual question answering. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2018, 40(10): 2413–2427. [doi: [10.1109/TPAMI.2017.2754246](https://doi.org/10.1109/TPAMI.2017.2754246)]
- [16] Yu Z, Yu J, Cui YH, Tao DC, Tian Q. Deep modular co-attention networks for visual question answering. In: *Proc. of the 2019 IEEE/CVF Conf. on Computer Vision and Pattern Recognition. Long Beach: IEEE*, 2019. 6274–6283. [doi: [10.1109/CVPR.2019.00644](https://doi.org/10.1109/CVPR.2019.00644)]
- [17] Song LY, Li JN, Liu J, Yang Y, Shang XQ, Sun MX. Answering knowledge-based visual questions via the exploration of question purpose. *Pattern Recognition*, 2023, 133: 109015. [doi: [10.1016/j.patcog.2022.109015](https://doi.org/10.1016/j.patcog.2022.109015)]
- [18] Liu H, Ji RR, Wu YJ, Hua G. Supervised matrix factorization for cross-modality hashing. In: *Proc. of the 25th Int'l Joint Conf. on Artificial Intelligence. New York: AAAI*, 2016. 1767–1773.
- [19] Liu H, Ji RR, Wu YJ, Huang FY, Zhang BC. Cross-modality binary code learning via fusion similarity hashing. In: *Proc. of the 2017 IEEE Conf. on Computer Vision and Pattern Recognition. Honolulu: IEEE*, 2017. 6345–6353. [doi: [10.1109/CVPR.2017.672](https://doi.org/10.1109/CVPR.2017.672)]
- [20] Shen HT, Liu LC, Yang Y, Xu X, Huang Z, Shen FM, Hong RC. Exploiting subspace relation in semantic labels for cross-modal

- hashing. *IEEE Trans. on Knowledge and Data Engineering*, 2021, 33(10): 3351–3365. [doi: [10.1109/TKDE.2020.2970050](https://doi.org/10.1109/TKDE.2020.2970050)]
- [21] Tu RC, Mao XL, Ma B, Hu Y, Yan T, Wei W, Huang HY. Deep cross-modal hashing with hashing functions and unified hash codes jointly learning. *IEEE Trans. on Knowledge and Data Engineering*, 2022, 34(2): 560–572. [doi: [10.1109/TKDE.2020.2987312](https://doi.org/10.1109/TKDE.2020.2987312)]
- [22] Song LY, Shang XQ, Yang C, Sun MX. Attribute-guided multiple instance hashing network for cross-modal zero-shot hashing. *IEEE Trans. on Multimedia*, 2022, 25: 5305–5318. [doi: [10.1109/TMM.2022.3190222](https://doi.org/10.1109/TMM.2022.3190222)]
- [23] Yang XY, Lyu YF, Tian T, Liu YF, Liu YD, Zhang X. Rumor detection on social media with graph structured adversarial learning. In: Proc. of the 29th Int'l Joint Conf. on Artificial Intelligence. Yokohama: IJCAI, 2020. 1417–1423.
- [24] Wei LW, Hu D, Zhou W, Yue ZJ, Hu SL. Towards propagation uncertainty: Edge-enhanced Bayesian graph convolutional networks for rumor detection. In: Proc. of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th Int'l Joint Conf. on Natural Language Processing (Vol. 1: Long Papers). ACL, 2021. 3845–3854. [doi: [10.18653/v1/2021.acl-long.297](https://doi.org/10.18653/v1/2021.acl-long.297)]
- [25] Gao L, Song LY, Liu J, Chen BL, Shang XQ. Topology imbalance and relation inauthenticity aware hierarchical graph attention networks for fake news detection. In: Proc. of the 29th Int'l Conf. on Computational Linguistics. Gyeongju: Int'l Committee on Computational Linguistics, 2022. 4687–4696.
- [26] Feng YF, You HX, Zhang ZZ, Ji RR, Gao Y. Hypergraph neural networks. In: Proc. of the 2019 AAAI Conf. on Artificial Intelligence. Honolulu: AAAI, 2019. 3558–3565. [doi: [10.1609/aaai.v33i01.33013558](https://doi.org/10.1609/aaai.v33i01.33013558)]
- [27] Wang X, Ji HY, Shi C, Wang B, Ye YF, Cui P, Yu PS. Heterogeneous graph attention network. In: Proc. of the 2019 World Wide Web Conf. San Francisco: ACM, 2019. 2022–2032. [doi: [10.1145/3308558.3313562](https://doi.org/10.1145/3308558.3313562)]
- [28] Hu ZN, Dong YX, Wang KS, Sun YZ. Heterogeneous graph transformer. In: Proc. of the 2020 Web Conf. Taipei: ACM, 2020. 2704–2710. [doi: [10.1145/3366423.3380027](https://doi.org/10.1145/3366423.3380027)]
- [29] Ding KZ, Wang JL, Li JD, Li DC, Liu H. Be more with less: Hypergraph attention networks for inductive text classification. In: Proc. of the 2020 Conf. on Empirical Methods in Natural Language Processing (EMNLP). ACL, 2020. 4927–4936. [doi: [10.18653/v1/2020.emnlp-main.399](https://doi.org/10.18653/v1/2020.emnlp-main.399)]
- [30] Liu J, Song LY, Wang GT, Shang XQ. Meta-HGT: Metapath-aware hypergraph Transformer for heterogeneous information network embedding. *Neural Networks*, 2023, 157: 65–76. [doi: [10.1016/j.neunet.2022.08.028](https://doi.org/10.1016/j.neunet.2022.08.028)]
- [31] Wang JN, Zhang S, Xiao YH, Song R. A review on graph neural network methods in financial applications. *Journal of Data Science*, 2022, 20(2): 111–134. [doi: [10.6339/22-JDS1047](https://doi.org/10.6339/22-JDS1047)]
- [32] Kurshan E, Shen HD. Graph computing for financial crime and fraud detection: Trends, challenges and outlook. *Int'l Journal of Semantic Computing*, 2020, 14(4): 565–589. [doi: [10.1142/S1793351X20300022](https://doi.org/10.1142/S1793351X20300022)]
- [33] Kurshan E, Shen HD, Yu HJ. Financial crime & fraud detection using graph computing: Application considerations & outlook. In: Proc. of the 2nd Int'l Conf. on Transdisciplinary AI (TransAI). Irvine: IEEE, 2020. 125–130. [doi: [10.1109/TransAI49837.2020.00029](https://doi.org/10.1109/TransAI49837.2020.00029)]
- [34] Rajput N, Singh K. Temporal graph learning for financial world: Algorithms, scalability, explainability & fairness. In: Proc. of the 28th ACM SIGKDD Conf. on Knowledge Discovery and Data Mining. Washington: ACM, 2022. 4818–4819. [doi: [10.1145/3534678.3542619](https://doi.org/10.1145/3534678.3542619)]
- [35] Skarding J, Gabrys B, Musial K. Foundations and modeling of dynamic networks using dynamic graph neural networks: A survey. *IEEE Access*, 2021, 9: 79143–79168. [doi: [10.1109/ACCESS.2021.3082932](https://doi.org/10.1109/ACCESS.2021.3082932)]
- [36] Zhu YC, Lyu FY, Hu CM, Chen X, Liu X. Encoder-decoder architecture for supervised dynamic graph learning: A survey. *arXiv:2203.10480*, 2022.
- [37] Kumar S, Zhang XK, Leskovec J. Predicting dynamic embedding trajectory in temporal interaction networks. In: Proc. of the 25th ACM SIGKDD Int'l Conf. on Knowledge Discovery & Data Mining. Anchorage: Association for Computing Machinery, 2019. 1269–1278. [doi: [10.1145/3292500.3330895](https://doi.org/10.1145/3292500.3330895)]
- [38] Trivedi R, Farajtabar M, Biswal P, Zha HY. DyRep: Learning representations over dynamic graphs. In: Proc. of the 7th Int'l Conf. on Learning Representations. New Orleans: ICLR, 2019.
- [39] Xu D, Ruan CW, Körpeoglu E, Kumar S, Achan K. Inductive representation learning on temporal graphs. In: Proc. of the 8th Int'l Conf. on Learning Representations. Addis Ababa: ICLR, 2020.
- [40] Rossi E, Chamberlain B, Frasca F, Eynard D, Monti F, Bronstein M. Temporal graph networks for deep learning on dynamic graphs. *arXiv:2006.10637*, 2020.
- [41] Tian S, Xiong T, Shi LL. Streaming dynamic graph neural networks for continuous-time temporal graph modeling. In: Proc. of the 2021 IEEE Int'l Conf. on Data Mining (ICDM). Auckland: IEEE, 2021. 1361–1366. [doi: [10.1109/ICDM51629.2021.00171](https://doi.org/10.1109/ICDM51629.2021.00171)]
- [42] Yang C, Wang CC, Lu YF, Gong XM, Shi C, Wang W, Zhang X. Few-shot link prediction in dynamic networks. In: Proc. of the 15th ACM Int'l Conf. on Web Search and Data Mining. New York: Association for Computing Machinery, 2022. 1245–1255. [doi: [10.1145/3534678.3542619](https://doi.org/10.1145/3534678.3542619)]

- 3488560.3498417]
- [43] Xia LH, Huang C, Xu Y, Dai P, Bo LF, Zhang XY, Chen TY. Spatial-temporal sequential hypergraph network for crime prediction with dynamic multiplex relation learning. In: Proc. of the 30th Int'l Joint Conf. on Artificial Intelligence. Montreal: IJCAI, 2021. 1631–1637.
 - [44] Li ZH, Huang C, Xia LH, Xu Y, Pei J. Spatial-temporal hypergraph self-supervised learning for crime prediction. In: Proc. of the 38th IEEE Int'l Conf. on Data Engineering (ICDE). Kuala Lumpur: IEEE, 2022. 2984–2996. [doi: [10.1109/ICDE53745.2022.00269](https://doi.org/10.1109/ICDE53745.2022.00269)]
 - [45] Zhang Y, Dong XY, Shang LY, Zhang DL, Wang D. A multi-modal graph neural network approach to traffic risk forecasting in smart urban sensing. In: Proc. of the 17th Annual IEEE Int'l Conf. on Sensing, Communication, and Networking (SECON). Como: IEEE, 2020. 1–9. [doi: [10.1109/SECON48991.2020.9158447](https://doi.org/10.1109/SECON48991.2020.9158447)]
 - [46] Wang DX, Zhang ZQ, Zhou J, Cui P, Fang JL, Jia QH, Fang YM, Qi Y. Temporal-aware graph neural network for credit risk prediction. In: Proc. of the 2021 SIAM Int'l Conf. on Data Mining. SDM, 2021. 702–710. [doi: [10.1137/1.9781611976700.79](https://doi.org/10.1137/1.9781611976700.79)]
 - [47] Zhang JF, Kuo AT, Zhao JN, Wen QL, Winstanley EL, Zhang CX, Ye YF. Rx-refill graph neural network to reduce drug overprescribing risks (extended abstract). In: Proc. of the 31st Int'l Joint Conf. on Artificial Intelligence. Vienna: IJCAI, 2022. 5379–5383.
 - [48] Yu B, Yin HT, Zhu ZX. Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting. In: Proc. of the 27th Int'l Joint Conf. on Artificial Intelligence. Stockholm: AAAI, 2018. 3634–3640.
 - [49] 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)]
 - [50] Wang XY, Ma Y, Wang YQ, Jin W, Wang X, Tang JL, Jia CY, Yu J. Traffic flow prediction via spatial temporal graph neural network. In: Proc. of the 2020 Web Conf. Taipei: ACM, 2020. 1082–1092. [doi: [10.1145/3366423.3380186](https://doi.org/10.1145/3366423.3380186)]
 - [51] Cui ZY, Henrickson K, Ke RM, Wang YH. Traffic graph convolutional recurrent neural network: A deep learning framework for network-scale traffic learning and forecasting. IEEE Trans. on Intelligent Transportation Systems, 2020, 21(11): 4883–4894. [doi: [10.1109/TITS.2019.2950416](https://doi.org/10.1109/TITS.2019.2950416)]
 - [52] Xu MX, Dai WR, Liu CM, Gao X, Lin WY, Qi GJ, Xiong HK. Spatial-temporal transformer networks for traffic flow forecasting. arXiv:2001.02908, 2021.
 - [53] Li H, Zhang SY, Li XJ, Su LC, Huang HJ, Jin D, Chen LH, Huang JB, Yoo J. DetectorNet: Transformer-enhanced spatial temporal graph neural network for traffic prediction. In: Proc. of the 29th Int'l Conf. on Advances in Geographic Information Systems. Beijing: Association for Computing Machinery, 2021. 133–136. [doi: [10.1145/3474717.3483920](https://doi.org/10.1145/3474717.3483920)]
 - [54] Lan SY, Ma YT, Huang WK, Wang WW, Yang HY, Li P. DSTAGNN: Dynamic spatial-temporal aware graph neural network for traffic flow forecasting. In: Proc. of the 39th Int'l Conf. on Machine Learning. Baltimore: ICML, 2022. 11906–11917.
 - [55] Yan SJ, Xiong YJ, Lin DH. Spatial temporal graph convolutional networks for skeleton-based action recognition. In: Proc. of the 32nd AAAI Conf. on Artificial Intelligence. New Orleans: AAAI, 2018. 912. [doi: [10.1609/aaai.v32i1.12328](https://doi.org/10.1609/aaai.v32i1.12328)]
 - [56] Yang D, Li MM, Fu H, Fan JC, Leung H. Unifying graph embedding features with graph convolutional networks for skeleton-based action recognition. arXiv:2003.03007, 2020.
 - [57] Hao XK, Li J, Guo YC, Jiang T, Yu M. Hypergraph neural network for skeleton-based action recognition. IEEE Trans. on Image Processing, 2021, 30: 2263–2275. [doi: [10.1109/TIP.2021.3051495](https://doi.org/10.1109/TIP.2021.3051495)]
 - [58] Plizzari C, Cannici M, Matteucci M. Spatial temporal transformer network for skeleton-based action recognition. In: Proc. of the 2021 Int'l Conf. on Pattern Recognition. Cham: Springer, 2021. 694–701. [doi: [10.1007/978-3-030-68796-0_50](https://doi.org/10.1007/978-3-030-68796-0_50)]
 - [59] Perozzi B, Al-Rfou R, Skiena S. DeepWalk: Online learning of social representations. In: Proc. of the 20th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. New York: Association for Computing Machinery, 2014. 701–710. [doi: [10.1145/2623330.2623732](https://doi.org/10.1145/2623330.2623732)]
 - [60] Grover A, Leskovec J. node2vec: Scalable feature learning for networks. In: Proc. of the 22nd ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. San Francisco: Association for Computing Machinery, 2016. 855–864. [doi: [10.1145/2939672.2939754](https://doi.org/10.1145/2939672.2939754)]
 - [61] Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. arXiv:1609.02907, 2017.
 - [62] Mnih V, Heess N, Graves A, Kavukcuoglu K. Recurrent models of visual attention. In: Proc. of the 27th Int'l Conf. on Neural Information Processing Systems. Montreal: MIT Press, 2014. 2204–2212.
 - [63] Feng FL, He XN, Wang X, Luo C, Liu YQ, Chua TS. Temporal relational ranking for stock prediction. ACM Trans. on Information Systems, 2019, 37(2): 27. [doi: [10.1145/3309547](https://doi.org/10.1145/3309547)]
 - [64] Ying XT, Xu C, Gao JL, Wang JX, Li Z. Time-aware graph relational attention network for stock recommendation. In: Proc. of the 29th ACM Int'l Conf. on Information & Knowledge Management. Ireland: ACM, 2020. 2281–2284. [doi: [10.1145/3340531.3412160](https://doi.org/10.1145/3340531.3412160)]
 - [65] Hou XR, Wang K, Zhong C, Wei Z. ST-Trader: A spatial-temporal deep neural network for modeling stock market movement.

- IEEE/CAA Journal of Automatica Sinica, 2021, 8(5): 1015–1024. [doi: [10.1109/JAS.2021.1003976](https://doi.org/10.1109/JAS.2021.1003976)]
- [66] Wu JR, Xu K, Chen XY, Li SZ, Zhao JC. Price graphs: Utilizing the structural information of financial time series for stock prediction. *Information Sciences*, 2022, 588: 405–424. [doi: [10.1016/j.ins.2021.12.089](https://doi.org/10.1016/j.ins.2021.12.089)]
- [67] Tian H, Zheng XL, Zhao K, Liu MW, Zeng DD. Inductive representation learning on dynamic stock co-movement graphs for stock predictions. *INFORMS Journal on Computing*, 2022, 34(4): 1940–1957. [doi: [10.1287/ijoc.2022.1172](https://doi.org/10.1287/ijoc.2022.1172)]
- [68] Cheng DW, Zhang YY, Yang FZ, Tu Y, Niu ZB, Zhang LQ. A dynamic default prediction framework for networked-guarantee loans. In: *Proc. of the 28th ACM Int'l Conf. on Information and Knowledge Management*. Beijing: ACM, 2019. 2547–2555. [doi: [10.1145/3357384.3357804](https://doi.org/10.1145/3357384.3357804)]
- [69] Cheng DW, Niu ZB, Zhang LQ. Delinquent events prediction in temporal networked-guarantee loans. *IEEE Trans. on Neural Networks and Learning Systems*, 2023, 34(4): 1692–1704. [doi: [10.1109/TNNLS.2020.3027346](https://doi.org/10.1109/TNNLS.2020.3027346)]
- [70] Yang S, Zhang ZQ, Zhou J, Wang Y, Sun W, Zhong XY, Fang YM, Yu Q, Qi Y. Financial risk analysis for SMEs with graph-based supply chain mining. In: *Proc. of the 29th Int'l Joint Conf. on Artificial Intelligence*. Yokohama: IJCAI, 2021. 643.
- [71] Rao SX, Lanfranchi C, Zhang S, Han ZC, Zhang ZT, Min W, Cheng M, Shan YN, Zhao Y, Zhang C. Modelling graph dynamics in fraud detection with attention. *arXiv:2204.10614*, 2022.
- [72] Xia PF, Ni ZW, Xiao HW, Zhu XH, Peng P. A novel spatiotemporal prediction approach based on graph convolution neural networks and long short-term memory for money laundering fraud. *Arabian Journal for Science and Engineering*, 2022, 47(2): 1921–1937. [doi: [10.1007/s13369-021-06116-2](https://doi.org/10.1007/s13369-021-06116-2)]
- [73] Alarab I, Prakoonwit S. Graph-based LSTM for anti-money laundering: Experimenting temporal graph convolutional network with bitcoin data. *Neural Processing Letters*, 2023, 55(1): 689–707. [doi: [10.1007/s11063-022-10904-8](https://doi.org/10.1007/s11063-022-10904-8)]
- [74] Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Computation*, 1997, 9(8): 1735–1780. [doi: [10.1162/neco.1997.9.8.1735](https://doi.org/10.1162/neco.1997.9.8.1735)]
- [75] Cui CR, Li XJ, Zhang CY, Guan WL, Wang M. Temporal-relational hypergraph tri-attention networks for stock trend prediction. *Pattern Recognition*, 2023, 143: 109759. [doi: [10.1016/j.patcog.2023.109759](https://doi.org/10.1016/j.patcog.2023.109759)]
- [76] Cheng DW, Niu ZB, Zhang YY. Contagious chain risk rating for networked-guarantee loans. In: *Proc. of the 26th ACM SIGKDD Int'l Conf. on Knowledge Discovery & Data Mining*. ACM, 2020. 2715–2723. [doi: [10.1145/3394486.3403322](https://doi.org/10.1145/3394486.3403322)]
- [77] Cheng DW, Wang XY, Zhang Y, Zhang LQ. Risk guarantee prediction in networked-loans. In: *Proc. of the 29th Int'l Joint Conf. on Artificial Intelligence*. Yokohama: IJCAI, 2021. 618.
- [78] Cho K, Van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, Bengio Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. In: *Proc. of the 2014 Conf. on Empirical Methods in Natural Language Processing (EMNLP)*. Doha: Association for Computational Linguistics, 2014. 1724–1734. [doi: [10.3115/v1/D14-1179](https://doi.org/10.3115/v1/D14-1179)]
- [79] Wang YY, Zheng QH, Ruan JF, Gao YD, Chen Y, Li XY, Dong B. T-EGAT: A temporal edge enhanced graph attention network for tax evasion detection. In: *Proc. of the 2020 IEEE Int'l Conf. on Big Data (Big Data)*. Atlanta: IEEE, 2020. 1410–1415. [doi: [10.1109/BigData50022.2020.9378157](https://doi.org/10.1109/BigData50022.2020.9378157)]
- [80] Ostmeier J, Cowell L. Machine learning on sequential data using a recurrent weighted average. *Neurocomputing*, 2019, 331: 281–288. [doi: [10.1016/j.neucom.2018.11.066](https://doi.org/10.1016/j.neucom.2018.11.066)]
- [81] Deng SM, Zhang NY, Zhang W, Chen JY, Pan JZ, Chen HJ. Knowledge-driven stock trend prediction and explanation via temporal convolutional network. In: *Companion Proc. of the 2019 World Wide Web Conf.* San Francisco: ACM, 2019. 678–685. [doi: [10.1145/3308560.3317701](https://doi.org/10.1145/3308560.3317701)]
- [82] Wu ZH, Pan SR, Long GD, Jiang J, Zhang CQ. Graph WaveNet for deep spatial-temporal graph modeling. In: *Proc. of the 28th Int'l Joint Conf. on Artificial Intelligence*. Macao: IJCAI, 2019. 1907–1913.
- [83] 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*. ACM, 2020. 753–763. [doi: [10.1145/3394486.3403118](https://doi.org/10.1145/3394486.3403118)]
- [84] Cao DF, Wang YJ, Duan JY, Zhang C, Zhu X, Huang CR, Tong YH, Xu BX, Bai J, Tong J, Zhang Q. Spectral temporal graph neural network for multivariate time-series forecasting. In: *Proc. of the 34th Int'l Conf. on Neural Information Processing Systems*. Vancouver: Curran Associates Inc., 2020. 1491.
- [85] Sawhney R, Agarwal S, Wadhwa A, Shah RR. Spatiotemporal hypergraph convolution network for stock movement forecasting. In: *Proc. of the 2020 IEEE Int'l Conf. on Data Mining (ICDM)*. Sorrento: IEEE, 2020. 482–491. [doi: [10.1109/ICDM50108.2020.00057](https://doi.org/10.1109/ICDM50108.2020.00057)]
- [86] Sawhney R, Agarwal S, Thakkar M, Malhotra V, Chava S. Time evolving spatio-temporal hypergraph for stock forecasting. In: *Proc. of the 15th ACM Int'l Conf. on Web Search and Data Mining*. New York: ACM, 2022.

- [87] Zhang X, You JL. A gated dilated causal convolution based encoder-decoder for network traffic forecasting. *IEEE Access*, 2020, 8: 6087–6097. [doi: [10.1109/ACCESS.2019.2963449](https://doi.org/10.1109/ACCESS.2019.2963449)]
- [88] Rao SX, Zhang S, Han Z, Zhang Z, Min W, Cheng M, Shan Y, Zhao Y, Zhang C. Suspicious massive registration detection via dynamic heterogeneous graph neural networks. *arXiv:2012.10831*, 2020.
- [89] Cheng DW, Wang XY, Zhang Y, Zhang LQ. Graph neural network for fraud detection via spatial-temporal attention. *IEEE Trans. on Knowledge and Data Engineering*, 2022, 34(8): 3800–3813. [doi: [10.1109/TKDE.2020.3025588](https://doi.org/10.1109/TKDE.2020.3025588)]
- [90] Xiang S, Cheng DW, Shang CC, Zhang Y, Liang YQ. Temporal and heterogeneous graph neural network for financial time series prediction. In: *Proc. of the 31st ACM Int'l Conf. on Information & Knowledge Management*. Atlanta: Association for Computing Machinery, 2022. 3584–3593. [doi: [10.1145/3511808.3557089](https://doi.org/10.1145/3511808.3557089)]
- [91] 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.
- [92] Wang XH, Lyu D, Li MJ, Xia Y, Yang Q, Wang XW, Wang XG, Cui P, Yang YP, Sun BW, Guo ZY. APAN: Asynchronous propagation attention network for real-time temporal graph embedding. In: *Proc. of the 2021 Int'l Conf. on Management of Data*. ACM, 2021. 2628–2638. [doi: [10.1145/3448016.3457564](https://doi.org/10.1145/3448016.3457564)]
- [93] Sawhney R, Agarwal S, Wadhwa A, Derr T, Shah RR. Stock selection via spatiotemporal hypergraph attention network: A learning to rank approach. In: *Proc. of the 35th AAAI Conf. on Artificial Intelligence*. AAAI, 2021. 497–504. [doi: [10.1609/aaai.v35i1.16127](https://doi.org/10.1609/aaai.v35i1.16127)]
- [94] Bacry E, Mastromatteo I, Muzy JF. Hawkes processes in finance. *arXiv:1502.04592v2*, 2015.
- [95] Lu MX, Han ZC, Rao SX, Zhang ZT, Zhao Y, Shan YN, Raghunathan R, Zhang C, Jiang JW. BRIGHT—Graph neural networks in real-time fraud detection. In: *Proc. of the 31st ACM Int'l Conf. on Information & Knowledge Management*. Atlanta: Association for Computing Machinery, 2022. 3342–3351. [doi: [10.1145/3511808.3557136](https://doi.org/10.1145/3511808.3557136)]
- [96] Reddy S, Poduval P, Chauhan AVS, Singh M, Verma S, Singh K, Bhowmik T. TeGraF: Temporal and graph based fraudulent transaction detection framework. In: *Proc. of the 2nd ACM Int'l Conf. on AI in Finance*. ACM, 2021. 18. [doi: [10.1145/3490354.3494383](https://doi.org/10.1145/3490354.3494383)]
- [97] Bello I. LambdaNetworks: Modeling long-range interactions without attention. In: *Proc. of the 9th Int'l Conf. on Learning Representations*. ICLR, 2021.
- [98] Daley DJ, Vere-Jones D. *An Introduction to the Theory of Point Processes: Vol. I: Elementary Theory and Methods*. New York: Springer, 2003. [doi: [10.1007/b97277](https://doi.org/10.1007/b97277)]
- [99] Assumpção HS, Souza F, Campos LL, De Castro Pires VT, De Almeida PML, Murai F. DELATOR: Money laundering detection via multi-task learning on large transaction graphs. In: *Proc. of the 2022 IEEE Int'l Conf. on Big Data*. Osaka: IEEE, 2022. 709–714. [doi: [10.1109/BigData55660.2022.10021010](https://doi.org/10.1109/BigData55660.2022.10021010)]
- [100] Lea C, Vidal R, Reiter A, Hager GD. Temporal convolutional networks: A unified approach to action segmentation. In: *Proc. of the 2016 European Conf. on Computer Vision*. Amsterdam: Springer, 2016. 47–54. [doi: [10.1007/978-3-319-49409-8_7](https://doi.org/10.1007/978-3-319-49409-8_7)]
- [101] Lea C, Flynn MD, Vidal R, Reiter A, Hager GD. Temporal convolutional networks for action segmentation and detection. In: *Proc. of the 2017 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*. Honolulu: IEEE, 2017. 1003–1012. [doi: [10.1109/CVPR.2017.113](https://doi.org/10.1109/CVPR.2017.113)]
- [102] Bai SJ, Kolter JZ, Koltun V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv:1803.01271*, 2018.
- [103] Box GEP, Pierce DA. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, 1970, 65(332): 1509–1526. [doi: [10.1080/01621459.1970.10481180](https://doi.org/10.1080/01621459.1970.10481180)]
- [104] Bollerslev T. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 1986, 31(3): 307–327. [doi: [10.1016/0304-4076\(86\)90063-1](https://doi.org/10.1016/0304-4076(86)90063-1)]
- [105] Cortes C, Vapnik V. Support-vector networks. *Machine Learning*, 1995, 20(3): 273–297. [doi: [10.1023/A:1022627411411](https://doi.org/10.1023/A:1022627411411)]
- [106] Breiman L. Random forests. *Machine Learning*, 2001, 45(1): 5–32. [doi: [10.1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324)]
- [107] Degutis A, Novickytė L. THE efficient market hypothesis: A critical review of literature and methodology. *Ekonomika*, 2014, 93(2): 7–23. [doi: [10.15388/Ekon.2014.2.3549](https://doi.org/10.15388/Ekon.2014.2.3549)]
- [108] Lacasa L, Luque B, Ballesteros F, Luque J, Nuño JC. From time series to complex networks: The visibility graph. *Proc. of the National Academy of Sciences of the United States of America*, 2008, 105(13): 4972–4975. [doi: [10.1073/pnas.0709247105](https://doi.org/10.1073/pnas.0709247105)]
- [109] Ribeiro LFR, Saverese PHP, Figueiredo DR. struc2vec: Learning node representations from structural identity. In: *Proc. of the 23rd ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining*. Halifax: Association for Computing Machinery, 2017. 385–394. [doi: [10.1145/3097983.3098061](https://doi.org/10.1145/3097983.3098061)]

- [110] Long W, Lu ZC, Cui LX. Deep learning-based feature engineering for stock price movement prediction. Knowledge-based Systems, 2019, 164: 163–173. [doi: 10.1016/j.knosys.2018.10.034]
- [111] Li W, Bao RH, Harimoto K, Chen DL, Xu JJ, Su Q. Modeling the stock relation with graph network for overnight stock movement prediction. In: Proc. of the 29th Int'l Joint Conf. on Artificial Intelligence. Yokohama: IJCAI, 2020. 626.
- [112] Xiang S, Wen D, Cheng DW, Zhang Y, Qin L, Qian ZP, Lin XM. General graph generators: Experiments, analyses, and improvements. The VLDB Journal, 2022, 31(5): 897–925. [doi: 10.1007/s00778-021-00701-5]
- [113] Pillay K, Moodley D. Exploring graph neural networks for stock market prediction on the JSE. In: Proc. of the 2nd Southern African Conf. for Artificial Intelligence Research. Durban: Springer, 2022. 95–110. [doi: 10.1007/978-3-030-95070-5_7]
- [114] Cheng DW, Niu ZB, Li J, Jiang CJ. Regulating systemic crises: Stemming the contagion risk in networked-loans through deep graph learning. IEEE Trans. on Knowledge and Data Engineering, 2023, 35(6): 6278–6289. [doi: 10.1109/TKDE.2022.3162339]
- [115] Huang XW, Yang Y, Wang Y, Wang CP, Zhang ZS, Xu JR, Chen L, Vazirgiannis M. DGraph: A large-scale financial dataset for graph anomaly detection. In: Proc. of the 36th Advances in Neural Information Processing Systems. New Orleans, 2022. 22765–22777.
- [116] Nakamoto S. Bitcoin: A peer-to-peer electronic cash system. 2008. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3440802
- [117] Weber M, Chen J, Suzumura T, Pareja A, Ma TF, Kanezashi H, Kaler T, Leiserson CE, Schardl TB. Scalable graph learning for anti-money laundering: A first look. arXiv:1812.00076, 2018.
- [118] Chen J, Ma TF, Xiao C. FastGCN: Fast learning with graph convolutional networks via importance sampling. In: Proc. of the 6th Int'l Conf. on Learning Representations. Vancouver: ICLR, 2018.
- [119] Weber M, Domeniconi G, Chen J, Weidele DKI, Bellei C, Robinson T, Leiserson CE. Anti-money laundering in bitcoin: Experimenting with graph convolutional networks for financial forensics. arXiv:1908.02591, 2019.
- [120] Zhao TX, Zhang X, Wang SH. GraphSMOTE: Imbalanced node classification on graphs with graph neural networks. In: Proc. of the 14th ACM Int'l Conf. on Web Search and Data Mining. ACM, 2021. 833–841. [doi: 10.1145/3437963.3441720]
- [121] Rao SX, Zhang S, Han ZC, Zhang ZT, Min W, Chen ZY, Shan YN, Zhao Y, Zhang C. xFraud: Explainable fraud transaction detection. Proc. of the VLDB Endowment, 2021, 15(3): 427–436. [doi: 10.14778/3494124.3494128]
- [122] Liu J, Shang XQ, Song LY, Tan YC. Progress of graph neural networks on complex graph mining. Ruan Jian Xue Bao/Journal of Software, 2022, 33(10): 3582–3618 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/6626.htm> [doi: 10.13328/j.cnki.jos.006626]

附中文参考文献:

- [122] 刘杰, 尚学群, 宋凌云, 谭亚聪. 图神经网络在复杂图挖掘上的研究进展. 软件学报, 2022, 33(10): 3582–3618. <http://www.jos.org.cn/1000-9825/6626.htm> [doi: 10.13328/j.cnki.jos.006626]



宋凌云(1988—), 男, 博士, 副教授, CCF 专业会员, 主要研究领域为计算机视觉, 基于金融大数据分析的风险预测, 基于教育大数据的认知诊断。



李战怀(1961—), 男, 博士, 教授, 博士生导师, CCF 高级会员, 主要研究领域为大数据管理技术, 海量信息存储系统。



马卓源(1999—), 男, 硕士生, 主要研究领域为时空图神经网络, 金融大数据。



尚学群(1973—), 女, 博士, 教授, 博士生导师, CCF 高级会员, 主要研究领域为数据挖掘, 机器学习, 生物信息学, 大数据。