

基于实时动态图联合学习框架的金融交易风控技术

周 俊^{1,2}, 曹月恬², 胡斌斌², 张志强², 陈超超¹

(1. 浙江大学计算机科学与技术学院, 浙江杭州 310027; 2. 蚂蚁科技集团股份有限公司机器智能部, 浙江杭州 310000)

摘要: 金融交易风险防控是金融风控平台最重要的能力之一. 近年来, 随着金融风控平台智能化需求的不断升级, 对其中相关应用算法的性能要求也水涨船高. 目前业界已完成了两代针对金融交易行为的表征学习框架的迭代升级. 第一代框架引入了金融交易活动参与者自身的历史行为序列, 利用序列模型学习其历史行为特征. 第二代框架通过一套实时大数据系统对资金流图进行建模, 根据业务专家预定义的业务规则计算出需要的实时特征, 并将其输入到后续的判别模型中. 相比第一代, 第二代框架引入了更多实时动态资金流图上的交互信息, 因而取得了不错的性能提升. 然而, 第二代框架在精细化、智能化和时序建模方面仍存在较大不足. 为了解决这些问题, 本文针对性地设计了第三代框架, 该框架通过动态图表征学习算法, 从实时资金流图的原始数据中直接进行表征学习, 以此规避了第二代框架中的诸多问题. 总的来说, 本文在时序信息建模和动态图框架层面均进行了创新性设计. 在时序信息建模层面, 利用了 C²GAT 模块 (连续时间和上下文感知的图注意力神经网络), 在动态多变的资金流图上快速地捕捉了高阶的结构化时序状态与信息. 在动态图框架层面, 开发了一套联合实时动态图表征框架——RULF, 该框架可以实时刻画出金融场景中多用户资金行为中存在的特定模式. 将金融场景中多角色联合行为和单角色独立行为进行了显式地解耦, 并将多个子图模块联合起来学习, 通过学习到更精准的行为表征, 从而更进一步地提高下游判别模型的性能. 本文将花呗套现交易识别——一个典型的金融交易风控场景为例, 介绍该框架在实际业务场景中的设计思想和实现细节.

关键词: 时序建模; 实时动态图; 图学习; 注意力机制; 深度学习; 系统框架

基金项目: 国家自然科学基金 (No.72192823, No.62172362)

中图分类号: TP183

文献标识码: A

文章编号: 0372-2112(2023)10-2801-11

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20220812

Real-Time Dynamic Graph Unified Learning Framework for Financial Transaction Risk Management

ZHOU Jun^{1,2}, CAO Yue-tian², HU Bin-bin², ZHANG Zhi-qiang², CHEN Chao-chao¹

(1. College of Computer Science and Technology, Zhejiang University, Hangzhou, Zhejiang 310027, China;

2. Department of Machine Intelligence, Ant Group, Hangzhou, Zhejiang 310000, China)

Abstract: In recent years, with the continuous escalation of demand in the intelligent financial platforms, the performance requirements for these relevant application algorithms in financial scenarios have also risen. At present, two generations of frameworks about financial role representation learning have been widely used in the industry. The first-generation framework introduced the unique historical sequence of financial roles, and used the sequence model to learn the historical behavior of the role. The second-generation framework put more emphasis on the interaction between roles, built a real-time dynamic graph system through capital flow, and directly obtained the required real-time features through graph calculation according to predefined business rules, and added them to the follow-up discriminant models. Compared with the first generation, it introduced more interactive information, resulting in a good performance improvement. However, the second-generation framework still has great limitations in terms of timeliness, generalization, and ease of use. In order to solve these problems, we design the third-generation framework which directly builds feature from the original real-time capital flow graph through the dynamic graph learning algorithm, avoiding many problems in the second generation. This paper mainly carries on the innovative design in temporal modeling and frame design. In terms of temporal modeling, we design the C²GAT to flexibly capture high-order structured temporal information on dynamic graphs. In terms of framework mod-

eling, we design a real-time dynamic graph framework—RULF, which can better capture and characterize the specific patterns existing in capital behavior in real time financial scenarios. We explicitly separate multi-role joint behavior and single-role independent behavior in financial scenarios, and jointly learn multiple subgraph modules to obtain accurate user representation and performance improvement. A typical interactive financial scenario will be used as a credit cashback example in this article to introduce our design ideas and implementation details in actual business scenarios.

Key words: time encoding; real time dynamic graph; graph learning; attention mechanism; deep learning; system framework

Foundation Item(s): National Natural Science Foundation of China (No.72192823, No.62172362)

1 引言

近年来,随着大数据系统数据收集、存储及处理能力的不断增强,数据的量级和丰富度也得到了不断扩展.金融交易风险判别框架也变得越来越高精度化、高智能化和高可用化^[1].金融交易风控框架经历了多代升级,升级过程中一个共同趋势就是采用更细粒度、更实时的数据形式,并针对该数据形式搭建适配的算法^[2,3].下面是两次比较重要的架构升级.

第一代框架^[4,5],引入了金融活动参与者的历史行为序列并用序列模型进行学习,但由于序列是独立的,对于涉及多个参与者的场景(如交易),此类框架缺乏多角色之间的互动视角,而该视角对很多金融风控模型的性能具有重要意义.第二代框架,将金融场景中多个参与者的历史行为构建为一张在线实时的图.算法按照业务专家预先定义好的业务规则,直接从图中计算出一些实时特征,并将这些特征输入后续的判别模型中.

目前第二代框架已在业界获得了大规模广泛应用^[6].我们将其作为本文的基线,并指出该基线框架中存在三个主要问题.

(1)精细化程度不足.图计算得到的特征通常是一些统计量,如过去一段时间的交易转账的总次数、总金额、平均金额等,无法精细地刻画资金事件中的细节信息.

(2)智能化程度不足.特征提取依赖业务专家的业务经验,人工干预程度过重,导致对人力资源消耗较大.同时,由于人对新欺诈规则的发现往往具有滞后性,人工设定规则计算特征的方式对新的欺诈模式往往效果不佳.

(3)时序性与拓扑性建模程度不足,这两类特性很难被业务规则所捕获,而这两个因素恰巧是金融风控场景中十分重要的信息维度.

作为一种直接从原始数据进行建模或表征学习的算法,深度学习从2012年开始取得了井喷式发展.尽管传统深度学习方法^[7]在欧氏空间数据的表征提取方面取得了巨大成功,但许多实际应用场景中的数据却属于非欧式空间数据,传统深度学习方法在此类数据

上的表现难以令人满意.例如不规则的图(graph)结构数据,图中的每个节点都可以有不同数量的邻居节点,导致常规的卷积这种在图像(image)上容易应用的算子,不再适合直接应用于图机器学习.此外,传统深度学习算法的一个核心假设是数据样本之间彼此独立.然而对于图数据来说,图中的每个数据样本(节点)通常会有边与图中其他数据样本(节点)相连,这些边上的信息表征了节点之间存在的相互依赖或关联关系.

因此,图神经网络(Graph Neural Network, GNN)作为一种主要针对非欧空间结构(图结构)的数据智能算法,近年来得到了广泛关注.文献[8]提出了第一代基于卷积的图神经网络 GCN(Graph Convolutional Network).其从基于频谱理论的卷积定理出发,在谱空间中定义了图卷积,此类方法在后来的也被统称为谱方法.但最初的谱方法具有时空复杂度较高的弊端,后续研究者开始大量采用基于空间的图聚合模式.该类方法通过不断改进聚合算子,引入多样化的注意力机制^[9,10],实现了对节点间的权重进行建模,从而进一步精准刻画了邻居对中心节点的影响.随着聚合算子的逐渐完善,很多研究人员开始将研究重点转向图领域中特定的技术领域或子方向.在子图采样领域,文献[11]通过引入bandit算法提高了学习过程中的子图采样效率和质量.在图增强领域,文献[12]通过施加全局正则增加了图学习的鲁棒性.在跨域学习领域,文献[13]构建了基于图的迁移学习模型,缓解了模型难以学习稀疏特征节点表征的问题.文献[14]探索了图在对抗场景中的应用.文献[15]则针对分布偏移,一种图训练中较为常见的问题,提出了良好的解决方案.

除了上述基于技术领域的研究外,近年来还有大量研究的重点是关注基于特定业务模式下的特定图数据的.该类数据的一个特点是具有独特的形式、更细粒度的原始信息,研究者往往需要设计更定制化的算法对其进行精细化地建模.关于图神经网络的学习算法和技术落地方案,近年来已经在众多业务场景中取得了实际应用,并取得了不错的进展.文献[16]设计了基于时空图的一种算法落地方案.文献[17]基于金融供应链场景中的数据特殊性设计了高效的算法.文献[18]针对社交推荐场景中图数据的异构性设计了对应

的条件注意力图模型,并取得了良好的效果.文献[19]针对智能对话场景中标签图谱数据设计了定制化的图学习算法.文献[20]通过对问答系统中原始长句进行词性分类、命名实体识别、分词、关键词聚合等一系列预处理步骤,构建了基于语义块的知识图谱数据结构,利用图模型对其进行语义学习以建模关键词的表征.文献[21]使用GCN构建SDN(Software-Defined Networking)路由性能预测模型.相较之前的基线方法,通过对SDN信息网络中的原始数据进行直接学习,为了有效地进行后续的路由策略建设,图模型能更好地捕捉到网络拓扑中物理链路和路由策略路径之间错综复杂的关系.这样就能快速准确地预测给定拓扑结构、路由策略和流量矩阵下的性能指标,如延迟和抖动.因此,类似上述已被证明有效的先例方法,本文提出的框架亦采用直接从原始数据构造图结构数据,并以此为基础设计图学习算法进行建模的方式.

在智能化方面,本文采用了端到端的学习模式,将多个模块进行联合建模.不仅替代了人工干预,极大降低了开发和维护成本,并且具有较好的可扩展性.模型上线后,通过加入在线学习模块^[22],模型还可以及时捕捉到灰黑产欺诈行为变化的新模式,从而进一步优化和提升模型性能.

在时序拓扑建模^[23-25]方面,本文希望从实时动态图中学习到关于金融参与者的良好时序拓扑表征.在时间信息建模上,国内外主要有三类比较先进的时间序列表征学习方法:

(1)基于随机游走的方法^[26,27]直接将时间信息融合到随机游走序列中,利用传统的skip-gram技术得到动态表征.然而,这些方法并没有考虑到节点和边上所包含的大量信息,且不支持归纳学习.

(2)基于时间片的方法^[1,16]首先在几个时间片上训练表征,然后利用神经网络结合不同时间片上的表征,得到最终的节点表征.这种技术忽略了每个时间片内存在的动态性,在某些情况下,危及节点关系随时间演化的特性.

(3)基于连续时间的方法^[28,29]尝试在连续的时间上学习时间序列或时序图的表征.然而,这些技术中的大多数仍然无法支持归纳学习.更关键的是,它们进行时间编码时并没有考虑到时序图上的节点之间的相互作用或时间维度上的上下文信息及其影响.

综合以上思路,本文提出了实时动态图联合学习框架——RULF(Real-time dynamic graph Unified Learning Framework).该框架较好地解决了上一代基线框架中存在的几个问题,同时还进一步提升了金融欺诈行为的检测性能.本框架的主要贡献点在于:

(1)针对特定的金融场景数据,以更细粒度的视角

对其中的数据进行了精准刻画.并在最后的实验模块中,验证了通过对精细化数据的建模,相比使用人工特征这种基于统计式的粗粒度数据,性能/效率/扩展性取得了较大提升.从而启发更多的研究者贡献和补充类似的有价值的刻画算法,到同类型的风险防控模式中,以满足他们场景的具体要求.

(2)设计了基于端到端的多模块联合优化方案,并将其丝滑地嵌入大型交易风控部署系统中,提供了一个很好的将图聚合能力与风控结合起来的全链路架构,以及在工业级场景中的一些算法优化经验.值得注意的是,在实际运行过程中,我们的框架可以在极高并发压力下保持良好的运行性能和鲁棒性.

(3)利用连续时间与上下文感知的图注意力网络(C²GAT)^[30]作为基础组件进行堆叠,来学习动态的资金流图中的节点表征,这是首个在金融交易风控-大规模资金流图中应用动态图技术进行建模的工作.为了充分挖掘金融交易参与者自身所处交易环境和不同参与者之间交易背景的拓扑结构,本文分别对买卖家中心子图、买卖家间路径子图进行表征.最终本文提出的RULF框架从时序变化信息与空间拓扑信息两个维度全面地建模了金融交易参与者的行为.通过详细的阐述与对比,使交易风控潜在的动态模式能够被更充分研究.

为了清晰地描述框架的实现细节,本文首先介绍了金融交易风控中的一个经典场景——蚂蚁花呗套现交易识别场景,以及该场景所对应的实时动态图的存储形式.接着,进一步阐述了实时动态图联合学习框架中关于算法框架和系统架构方面的实现细节,并指出了该方案中对套现交易识别这一场景的针对性设计.除此之外,本文还使用了C²GAT时序信息表征算法作为基础组件,对时序信息这一重要维度进行了精细化刻画和表征.最后,在公开的学术和工业数据集上进行了广泛的实验,以证明该框架在现实世界的应用场景中比基线架构更有优势,并证实所提算法的性能可以超过目前的学术最优模型.

2 场景和术语介绍

2.1 蚂蚁花呗套现交易识别场景

蚂蚁花呗,是一种基于用户信用的消费信贷服务,申请开通后,用户将获得一定的消费额度,用户在消费时,可以预支蚂蚁花呗的额度,享受“先消费,后付款”的购物体验.但在实际应用中,存在一小部分用户,其使用花呗服务的目的不是为了消费,而是将信贷额度套取出来留作他用.他们实际上并未产生一笔真正的消费支出,因此该行为被称作“套现”.出于该类场景的通用性、合理分配资金、防控金融风险等多

方面因素的考虑,金融从业者需要将这样的套现交易识别出来并进行拦截.从行业经验来看,一笔交易若是套现交易,其中涉及的买/卖双方的短期资金流(转账和交易)中通常会存在一些异常模式.比如常见的套现模式有:

- (1) 买卖家近期已频繁和其他“套现黑户”做过类似的交易.
- (2) 买卖家各自的资金流序列在时间、金额上存在异常.
- (3) 买卖家之间存在资金回流闭环.
- (4) 买卖家之间的资金流链路上存在疑似套现中介的“中间人”.

因此,需在 RULF 框架的设计过程中充分考虑上述模式,才能取得良好的算法性能.

2.2 实时动态图

首先对实时动态图进行简单的定义.本文语境中使用的实时动态图,指一套可支持实时增删改查的在线图数据库系统.其中,对图中点、边的增删改的感知延迟控制在分钟级以内.在查询方面,能支持毫秒级延迟的实时查询,并支持复杂的子图查询语法和检索方式,以便算法可以自定义子图构造逻辑.在上述场景中,实时动态图中主要存储的是资金事件,如转账和交易等,事件上带有金额、平台、时间戳等信息,其中的时间戳记录了该事件产生的具体时间点.在实践中,由于点、边数量均在亿级以上,对于如此大规模的数据量,需要将事件(即边)的生命周期设置为 k 天,即某个事件在发生的 k 天后,会被从图数据中删除, k 取值一般不会太大.此外,值得注意的是,在金融交易风控场景,目标交易发生之前一段较短时间(如几分钟到几个小时)内交易参与者的行为对该笔交易是否涉嫌欺诈有非常重要的判断意义.同时考虑到系统性能与消耗资源的问题,在实时的欺诈交易识别场景中,我们只对近期的用户行为通过图数据的方式进行建模,而更长时间内的用户历史行为,则主要通过一些离线用户画像进行建模.在实际场景的实验也证明了这样建模可以实现识别精度与系统性能较好的折中.

图 1 是一个简单的实时动态图示意,显示了从 3 月 5 日 11 点 58 分到 3 月 5 日 12 点之间三个整点分钟时刻对应的实时动态图快照.其中,蓝色的边代表交易,绿色的边代表转账.如图所示,在 3 月 5 日 11 点 59 分这个时刻,GB 这条边因超过 3 天的生命周期而被删除,此后的快照将无法再观察到这条边.同样,在 3 月 5 日 11 点 59 分这个时刻,因 F 节点与图中原节点 B 发生了一笔新的交易,而新增了一条连接 F 和 B 的交易边.

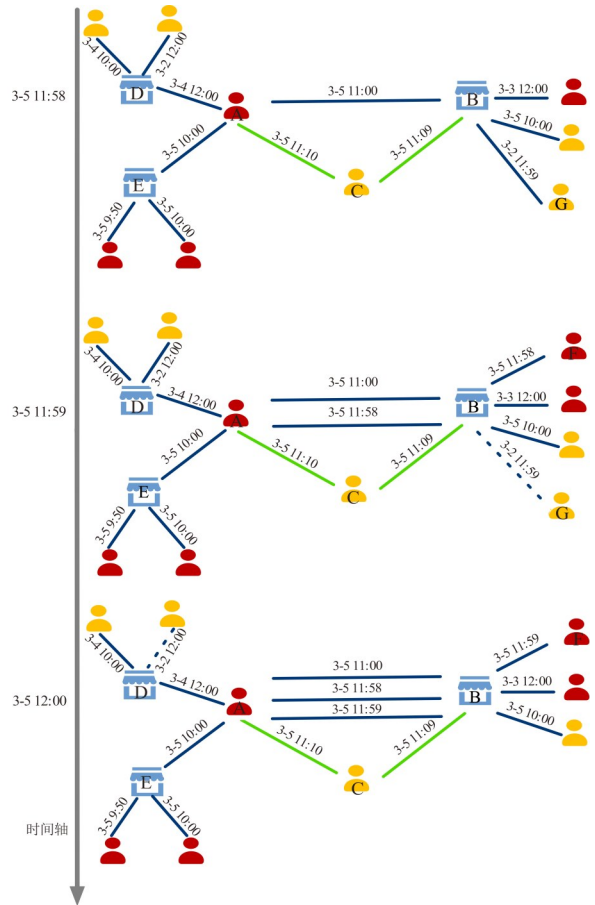


图 1 实时动态图示意

3 联合实时图表征学习框架

本文提出的 RULF 框架主要包含三大模块,分别是中心子图聚合模块、路径子图聚合模块和联合预测模块.算法整体框架如图 2 所示,系统架构图如图 3 所示.下面进一步介绍具体的实现细节.

3.1 构图逻辑

以交易样本发生的时刻为例,从实时动态图系统中取出对应时刻的图快照.由于只关心当前交易样本涉及的买卖家信息,所以需要从全量图中构造出判断套现行为所需的子图,以便后续作为动态图模型的输入.考虑到框架将最终运行在高并发系统中,同时结合套现交易识别场景中本身的一些业务特性,我们在构造子图之时进行了针对性设计,并在规模与性能之间进行了适当权衡.

(1) 规模.构造的子图需包含足够的信息,且这部分信息可以刻画出前文所描述的几种经典套现模式.不能做过度的裁剪.

(2) 性能.线上实际交易系统最高可达十亿级/天.在如此巨大的并发流量下,过度复杂的构图逻辑导致耗时的少量增加,都容易造成线上系统出现性能问题

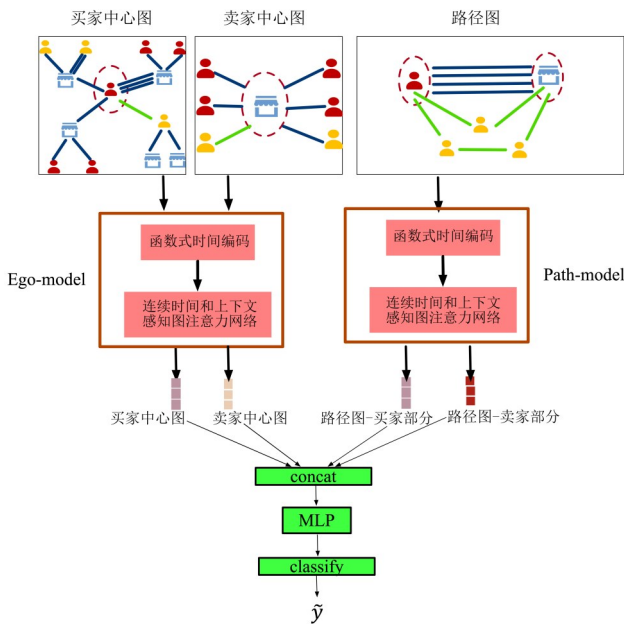


图2 联合学习框架

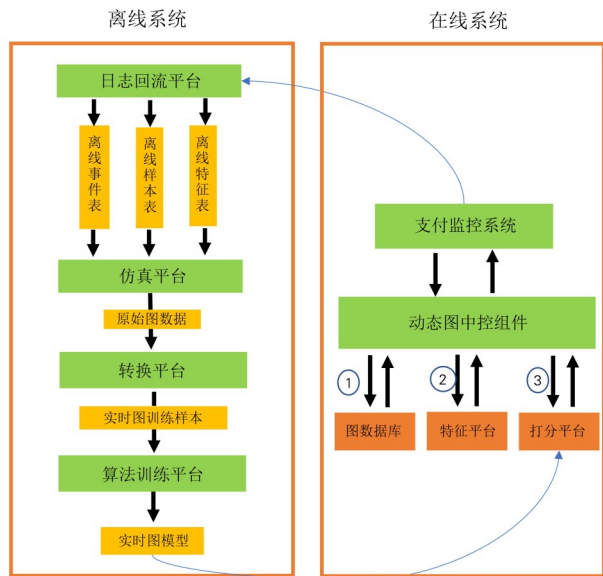


图3 系统架构

甚至崩溃,引发系统性风险。

按照链接预测的常规做法,首先以买卖家为中心构造带邻居的子图。通过对买卖家中心图中邻居聚集性的学习,可以很好地捕捉到套现模式 1、模式 2。根据业务分析,买家节点的度通常不大。所以对于买家我们选择保留尽量多的信息,少做裁剪。实践中选取了 2 跳邻居,每跳最多选取 30 条边。如果超过 30,会首先将边按时间排序,然后取最近的 30 条。该操作的逻辑是,我们认为新近发生的行为比远期的行为更重要。通过按时间先后顺序进行采样,保证了采样后的子图留下了买家最新近发生的行为事件集合。买家中心图如图 4 所示。

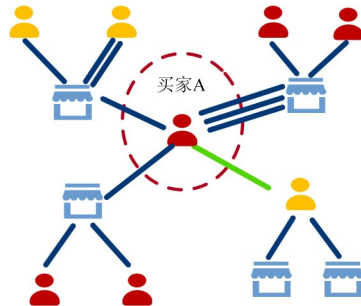


图4 买家中心图样例

由于不少卖家会在短期内产生数万级甚至更大规模的交易量,考虑到系统可能存在的性能瓶颈,以卖家为中心的邻居图必须进行大幅采样。实践中仅仅保留了卖家的一跳邻居,并根据预先计算出的套现事件中金额分布特点,以金额为尺度对卖家的边进行采样。通过这种启发式的做法,最大限度保证了采样后留下的边是对套现交易识别有贡献的事件,并能够让系统快速做出响应。卖家中心图如图 5 所示。

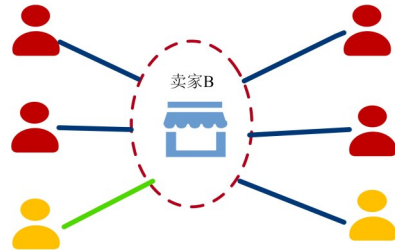


图5 卖家中心图样例

套现模式 3、模式 4 属于买卖双方之间的资金交互模式,因此为了充分地识别套现交易,需要在图中保留买卖双方之间的资金交互数据,即“互动视角”。但由于进行了采样,交易对手方的信息通常无法被涵盖在买卖双方各自的中心子图中。为了解决这个问题,本文在上述两种子图之外,还单独构造了买卖双方之间的连通路程子图。如图 6 所示。在构图时选择从买家出发,寻找从买家到卖家的连通路程,在全图中进行广度优先搜索,并至多扩展至 3 跳,即最长可选出“买家-中间人-中间人-卖家”这样的路径。该子图通过显示地提取买卖双方之间的资金交互数据,方便模型针对性地学习套现中经典的模式 3、模式 4。

3.2 联合学习算法框架

框架将上述 3 种子图联合起来进行同时学习。其中买卖家的中心子图目标都是刻画买卖家邻域的聚集性特征,因而在模型中共享了同一套聚合参数。路径子图因侧重学习买卖双方交互中存在的资金流模式,和中心子图的学习目标有较大差异,因而使用了独立的参数模块。通过 2 个图聚合模型,得到 4 个关于买卖家的表征之后(即买家的聚集性表征、卖家的聚集性表征、

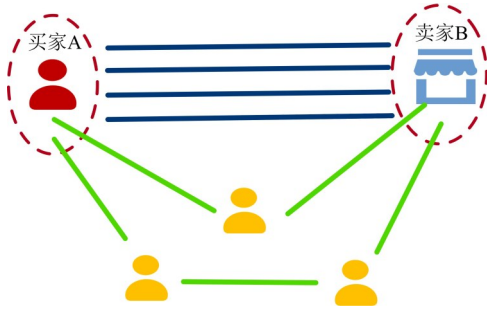


图6 路径子图样例

买家的交互表征、卖家的交互表征),这些表征最终被拼接起来送入判别模型,得到最终的预测结果。

3.3 系统架构

系统架构按运行环境可分为离线和在线系统两大部分,涉及的主要模块如图3所示(核心算法主要实现在“仿真/转换平台”与“算法训练平台”中)。离线系统通过接收线上回流的日志数据,按事件流、样本流和特征流进行整理,输入仿真回溯平台中构造出原始图数据,并将其转换为训练样本后利用图学习系统^[31]进行训练。在线系统则通过动态图中控组件实现与其他子模块之间的数据传输和通信调度,涉及的主要流程在图中已标示出,依次为:(1)从图数据库中查询子图结构;(2)从特征平台查询子图中的节点/边特征;(3)将①②两步查询结果进行整合后送往在线打分平台进行打分。

$$t \mapsto \Phi^M(t) := \left[\sqrt{c_1} \phi_1(t), \sqrt{c_2} \phi_2(t), \dots \right] \quad (1)$$

经验上来说,时序模式可以被一系列周期核函数刻画,参照文献[28]介绍的定理,以 ω 为频率的映射函数 $\Phi(\cdot)$ 可以进一步被形式化为

$$t \mapsto \Phi_\omega^M(t) := \left[\sqrt{c_1} \cos\left(\frac{j\pi t}{\omega}\right), \sqrt{c_2} \sin\left(\frac{j\pi t}{\omega}\right), \dots \right] \quad (2)$$

这种傅里叶级数形式的表示提供了更好的截断特性,因为被截断的 d 维映射函数 $\Phi_{\omega,d}^M(t)$ 可以近似原始的无穷维映射。随后,将 k 个周期(即 $\{\omega_1, \omega_2, \dots, \omega_k\}$)的映射函数拼接,形成最终的时间编码为

$$t \mapsto \Phi_d^M(t) := [\Phi_{\omega_1,d}^M(t) \parallel \dots \parallel \Phi_{\omega_k,d}^M(t)] \quad (3)$$

值得注意的是,这种时间编码是独立于节点的。换句话说,由于时间编码与节点无关,任何两个节点在同一时间间隔下都会有相同的时间编码。但在具体的应

$$v, t \mapsto \Phi_\omega^M(v, t) := \left[\sqrt{c_1(v)} \cos\left(\frac{j\pi t}{\omega}\right), \sqrt{c_2(v)} \sin\left(\frac{j\pi t}{\omega}\right), \dots \right] \quad (4)$$

$$v, t \mapsto \Phi_d^M(v, t) := [\Phi_{\omega_1,d}^M(v, t) \parallel \Phi_{\omega_2,d}^M(v, t) \parallel \dots \parallel \Phi_{\omega_k,d}^M(v, t)]$$

其中, $c_i(v): \mathbb{R}^d \rightarrow \mathbb{R}$ ($i=1, 2, \dots$) 是一个映射函数的集合,使用节点属性作为输入计算傅里叶系数。多层感知器在实验中被用来实现 $c_i(v)$,因为它对复杂的交互作用具有卓越的建模能力,即 $c_i(v) = \text{MLP}(f_v)$,其中 $f_v \in \mathbb{R}^d$ 是节点 v 的属性特征向量。并且,通过强制感知器的最后一层为正数的输出,我们满足了Mercer理论的固有特性。

实践中,每天的活跃买卖家去重后共约3亿,每天的事件量级约10亿,最大回看窗口设置为3天。所以理论上,线上数据库最大需要同时存储9亿点,30亿边。

4 动态图算法

我们利用连续时间与上下文感知的图注意力深度网络(C²GAT)对图中的时序变化信息进行建模,该方法最早在基于POI的推荐任务中被用于建模用户对本地生活类商品的偏好在时间上的分布。本文将这项技术应用于金融欺诈交易识别场景,下文将介绍这项技术并详细分析其是如何适配金融欺诈交易识别场景。

4.1 时间编码

如前所述,交易、转账事件(对应为图6上的边)上的时间戳是金融场景中非常关键的信息。因此本文设计了一个时序编码器,将时序信息映射到一个高维向量空间,以更精细地刻画其中涵盖的时间和时序特性。本文首先进行了一个简单的参考系转换,将事件上的绝对时间戳转换为相对于当前样本的相对时间。在此基础上,受到Mercer理论的启发,为了学习上述时序核函数及其映射函数,该映射函数被定义如下:

用场景中,以套现交易识别场景为例,即使相同的交易时间,对判断不同的买卖家对套现,其影响应该是完全不同的。比如,某类特征的用户习惯傍晚使用软件,对此类用户来说发生套现在傍晚的概率较大,而对另一类用户来说则更可能是清晨进行套现。因此,这样节点无关的时间编码不适合用于后续图神经网络来进行表征学习。因此,给定一个特定的节点 v ,进一步将节点相关的时间编码定义为

为了在连续的时间上对节点的时序模式进行建模,我们使用Mercer理论对动态图上每个节点相关的时间间隔进行时间编码。如图2所示,我们使用连续时间和上下文感知图注意力神经网络(C²GAT)自适应地学习图上的时序信息,一方面衡量各种相邻节点在时间上的相关性和重要性,另一方面考虑了相邻节点在时间维度上的相互作用^[30]。具体地,在C²GAT的第 l 层,

在时间 t 给定目标节点 v_i 的情况下,为该节点的邻居节点集合 $\mathcal{N}_{v_i}(t) = \{v | t_v < t\}$ 创建一个注意力分布,以融合每个邻居节点的表征. 平移不变性存在于定义的时间核函数中,因此我们使用 $\{t - t_v\}_{v \in \mathcal{N}_{v_i}(t)}$. 在 t 时刻,该节点对即目标节点 v_i 及其任意一个邻居 $v \in \mathcal{N}_{v_i}(t)$ 之间的连续时间和上下文感知的注意力值可以由此确定:

$$\alpha_{v_i, v}(t) = \frac{Q_{v_i}(t) K_v^T(t)}{\sqrt{d}} \quad (5)$$

$$Q_{v_i}(t) = [h_{v_i}^{(l-1)}(t) \| 0 \| \Phi_d^M(0)] W_Q$$

$$K_v(t) = [\mathcal{F}^{(l-1)}(v, t, t_v) \| e_{v_i, v}(t) \| \Phi_d^M(t - t_v)] W_K$$

其中, W_Q 和 W_K 是两个映射矩阵,前者用来获取“Query”矩阵,而后者用来获取“Key”矩阵; $e_{v_i, v}(t)$ 为 t 时刻节点 v_i 和节点 v 的边特征向量; $h_{v_i}^{(l-1)}$ 为节点在 C^2GAT 第 $l-1$ 层的输出. 为了保证更稳定的训练,很明显,这种注意机制可以灵活地扩展为多头注意机制. 如上所述,目标节点和邻居节点的注意力值的一个关键因素是节点邻居之间在时间上的上下文信息及其相互影响. 为了明确描述这样的影响,我们设计了一个名为 $\mathcal{F}^{(l-1)}(v, t, t_v)$ 的上下文聚合函数来实现该功能. 对于一个目标节点 v_i , 任何一个邻居节点 v 及其对应的时间 t_v , 定义了以下的邻居集合: $S_{v_i, v}(t_v) = \{v' \in \mathcal{N}_{v_i}(t) | t_{v'} \leq t_v\}$. 更进一步地,我们使用了两种聚合函数来实现 $\mathcal{F}(\cdot)$:

(1) 为了赋予 C^2GAT 强大的表达能力,循环聚合器在序列上下文上应用复杂的 LSTM 结构.

$$\mathcal{F}_R = \text{LSTM}(h_{v_i}^{(l-1)}(t); v' \in S_{u, v}(t_v)) \quad (6)$$

(2) 为了使 C^2GAT 可以扩展到更大规模数据集上,卷积聚合器使用了深度卷积操作.

$$\mathcal{F}_C = \sum_{i=1}^{d-1} \sum_{j=0}^{|S_{u, v}(t_v)|} h_j^{(l-1)}(t) W_{j, i} \quad (7)$$

其中, $W \in \mathbb{R}^{|S_{u, v}(t_v)| \times d \times d}$ 是卷积核.

通过叠加 L 层前述的 C^2GAT 层,可以从更广、更深的维度来利用动态图中的高阶结构信息. 由此,每个节点在 t 时刻的节点表征被记为 $\{h_v^z(t) = h_v^{(L)}(t)\}_{v \in \mathcal{V}}$. 由于用户表征是这项工作的重点,我们将每个用户在时间 t 的表征重写为 $h_u^z(t)$.

4.3 模型学习

在套现交易识别场景中,本文将样本 D 形式化定义为四元组 $D = (u, v, t, y)$. 其中 u 表示买家; v 表示卖家; t 表示样本发生的时刻; y 即标签,在套现交易识别场景中表示该笔交易在后续追溯期里是否被确认为套现交易. 本文最终利用的数据集是 S_D , 由于涉及的用户和交易量很大,采用随机负采样来近似计算. 根据极大似然估计的策略,最终的损失函数为

$$\sum_{(u, v, t, y) \in S_D} C(y, p(y^* | u, v, t)) + \|\omega\|_2 \quad (8)$$

其中, $S_D = S_D^+ \cup S_D^-$, S_D^- 表示对负样本进行固定比例的下采样后得到的所有负样本集合, S_D^+ 为训练集中的所有正样本集合; $C(\cdot, \cdot)$ 表示交叉熵损失函数; $\|\omega\|_2$ 表示使用了 L2 正则项.

5 实验分析

在本节,我们构造了一系列实验来验证设计的框架的有效性.

5.1 公开可用的数据集上的有效性实验

首先,我们在 4 个可公开获取的数据集 (Reddit, Wikipedia, MOOC, LastFM) 上进行 (推导/归纳) 链接预测和节点分类任务,以确认框架中时序信息建模模块 (即 C^2GAT) 的有效性. 表 1 显示了 4 个数据集的完整统计信息.

表 1 4 个公开可用的数据集的统计信息

	节点数	边数	特征数
Reddit	11 000	672 477	172
Wikipedia	9 227	157 474	172
MOOC	7 145	411 749	4
LastFM	2 000	1 293 103	—

本文为这个实验选择了 8 个代表性的基准技术,大致分为 3 类.

(1) 基于深度递归网络的 2 种技术,包括 Jodie^[5] 和 Time-LSTM^[4].

(2) 2 种静态的深度图神经网络方法: GraphSAGE^[32] 和 GAT^[9].

(3) 6 种动态深度图神经网络方法 (Time 缩写为 T), 包括 GraphSAGE-T, GAT-T, CTDNE^[33], M²DNE^[34], GCRN^[35], T²GAT^[28].

在 (推导/归纳) 链接预测和节点分类任务上,本文在表 2 和表 3 中分别报告了基于 C^2GAT 的 RULF 框架与上述 8 个基准方法的性能对比实验结果. 需要注意的是, RULF 有 2 个版本,其中 RULF-C 表示基于卷积聚合器实现的 C^2GAT , 而 RULF-R 表示基于循环聚合器实现的 C^2GAT . 可以从表 2 和表 3 中得出如下 2 点重要推论.

(1) 框架中堆叠的 C^2GAT 模型在链接预测和节点分类任务上一致大幅超过其余基准模型,证明了该模型的有效. 这是因为 C^2GAT 以更有效的方式利用了高阶结构信息和局部序列的上下文信息来捕捉图上的动态变化. 而基于循环聚合器的实现要普遍优于基于卷积聚合器的实现.

(2) 在这些对比的技术当中,绝大部分情况下,动态图模型优于静态图模型,也优于深度序列模型. 这凸

显了动态性、高阶结构和节点/边属性对学习动态图表 征非常关键.

表 2 公开数据集上(Tr.:推导式/In.:归纳式/AP:平均精度)链接预测任务的性能比较

方法	Reddit		Wikipedia		MOOC		LastFM		
	Accuracy	AP	Accuracy	AP	Accuracy	AP	Accuracy	AP	
Tr.	Time-LSTM	0.702 5	0.715 7	0.562 5	0.564 8	0.560 1	0.567 3	0.510 3	0.521 6
	Jodie	0.908 8	0.974 2	0.835 4	0.929 3	0.782 2	0.774 6	0.621 1	0.650 5
	GraphSAGE	0.932 3	0.983 0	0.888 9	0.959 9	0.700 4	0.745 9	0.644 1	0.692 2
	GAT	0.931 7	0.983 3	0.880 7	0.953 9	0.673 2	0.721 7	0.654 8	0.680 0
	CTDNE	0.781 0	0.859 4	0.552 1	0.568 9	0.580 2	0.591 9	0.392 0	0.439 9
	M ² DNE	0.862 2	0.942 9	0.816 7	0.909 1	0.685 8	0.694 5	0.592 6	0.620 1
	GCRN	0.933 8	0.982 9	0.885 5	0.955 2	0.710 6	0.746 2	0.654 1	0.721 3
	GraphSAGE-T	0.930 3	0.982 3	0.896 9	0.964 8	0.756 6	0.786 8	0.679 1	0.776 5
	GAT-T	0.932 3	0.983 4	0.898 4	0.964 7	0.755 3	0.790 1	0.678 5	0.757 6
	TGAT	0.934 2	0.983 7	0.874 3	0.950 2	0.686 9	0.715 7	0.676 5	0.673 2
	RULF-C	0.934 8	0.984 5	0.906 1	0.971 4	0.790 7	0.852 1	0.702 1	0.792 2
RULF-R	0.935 5	0.984 5	0.906 1	0.971 4	0.790 8	0.861 4	0.717 1	0.812 9	
In.	GraphSAGE	0.900 1	0.965 0	0.862 7	0.944 2	0.697 3	0.736 5	—	—
	GAT	0.901 8	0.966 9	0.854 3	0.937 2	0.661 0	0.699 7	—	—
	GCRN	0.900 2	0.963 6	0.853 3	0.932 8	0.694 2	0.743 8	—	—
	GraphSAGE-T	0.899 1	0.965 0	0.876 1	0.956 6	0.764 1	0.797 6	—	—
	GAT-T	0.903 1	0.968 1	0.880 3	0.956 2	0.767 7	0.804 3	—	—
	TGAT	0.905 6	0.967 9	0.853 6	0.935 3	0.678 9	0.703 6	—	—
	RULF-C	0.906 4	0.968 2	0.883 1	0.960 2	0.794 4	0.840 3	—	—
RULF-R	0.906 9	0.968 2	0.883 6	0.960 5	0.796 2	0.844 2	—	—	

表 3 公开数据集上节点分类任务的性能比较(AUC)

方法	Reddit	Wikipedia	MOOC
Time-LSTM	0.630 5	0.777 3	0.693 5
Jodie	0.610 6	0.762 9	0.665 5
GraphSAGE	0.650 2	0.797 4	0.674 2
GAT	0.661 7	0.848 0	0.645 9
GCRN	0.674 3	0.857 5	0.669 5
GraphSAGE-T	0.659 4	0.851 5	0.677 9
GAT-T	0.674 3	0.850 8	0.669 4
TGAT	0.641 1	0.860 6	0.678 6
RULF-R	0.684 6	0.875 3	0.688 1

5.2 工业数据集上的实验

接下来,我们使用花呗套现交易识别这个实际业务场景中数据,来构造对应的数据集,并以此进一步评估本框架的性能和效果.具体来说,我们选取了蚂蚁花呗中两个具有代表性的细分卖家客群——淘宝和线下商家,并构造对应的工业数据集.通常来说,一笔交易在其发生一个月后,通过回溯规则、人工检验等方式,被打上是/否套现的标签,因此本文采用的数据集具有一定滞后性.实践中,我们从过去一年的线上真实交易数据回流日志中抽取训练/测试样本.同时,为防止业务变动(如大促销)引发的数据分布倾斜,将数据集中

交易样本涉及的时间跨度尽量拉长,以平滑可能存在的短期数据分布波动.同时,由于套现事件发生的低频性,我们在构造训练数据时对负样本(即非套现交易)进行了部分下采样,通过增大正样本的相对占比来保证模型的学习能力.如何将动态图建模技术和非均衡数据建模方法相结合,留作未来需要深入研究的方向之一.具体来说,训练集合的正负样本比例为1:20,测试集为1:600.最后,表4显示了创建的2个工业数据集的统计数字.

表 4 2个工业数据集的统计信息

参数	淘宝数据集	线上商家数据集
买家数	4 316 万	5 585 万
卖家数	3 024 万	242 万
训练样本数	2 633 万	3 345 万
验证样本数	20 万	20 万
测试样本数	4 217 万	5 585 万
数据集时间跨度	5 个月	6 个月
日均测试样本数	28 万	37 万

作为需要最终落地的工业应用框架,我们从以下2点出发论证我们框架的有效性:(1)算法性能优于上一代框架;(2)新框架系统性能可以满足大规模系统的并发要求,且具有更加智能化的特性.下面我们对这2点

进行进一步阐述。

算法性能方面,我们选择满足一定召回比例时对应的准确率作为算法性能的评估指标,这是业界通用评估方法^[2]。具体来说,P@20R 表示召回 20% 正样本时的准确率,P@40R 表示召回 40% 正样本时的准确率。可以看到,相比基线系统,实时动态图框架的识别性能都具有提升。同时,我们还设计了一系列消融实验,以定量地观察实时动态图框架中各模块的相对贡献。实验数据见表 5。

表 5 工业数据集的性能比较及消融实验(“-”:去掉)

算法	淘宝交易场景		码商交易场景	
	P@20R	P@40R	P@20R	P@40R
MLP+实时特征	67.2	43.7	46.3	30.3
实时动态图	84.7	56.8	55.7	35.9
实时动态图-中心子图	81.0	50.7	54.5	33.1
实时动态图-路径子图	78.4	50.0	53.2	31.8
实时动态图-边特征	65.3	20.1	40.3	20.1
实时动态图-C ² GAT	80.1	54.5	53.8	33.2

从表 5 的前 2 行可以看出,相比基线框架,实时动态图算法在多指标、多数据集下的性能均有明显提升。同时,在套现交易识别这项实际业务应用中,业务更加看重准确率更高的头部正样本,即表 5 中所示的 P@20R 指标。可以看出,本文算法在准确抓取头部正样本时的能力更强,相比基线可以带来 10%~17% 左右的提升。

表 5 的后 4 行记录了关于消融实验的实验结果。通过只使用部分构图子模块进行训练,可以看出,中心子图和路径子图呈现出信息互补的特性,去掉其中任一模块,性能都会有明显损失。同时,我们还尝试了去掉子图中边上的特征,只保留图结构和连接关系,模型效果产生了大幅下降,由此可见图模型收益主要来源于动态图上的事件原始特征,从而进一步论证了直接从动态图原始数据粒度进行学习的必要性。此外,我们还尝试了去掉本文所设计的堆叠的时间编码模块 C²GAT,性能也出现了明显的下降,说明了时序信息建模通过较好地刻画场景中的时序信息,最终辅助判别模型更好地预测。

在系统性能方面,实时动态图框架在支付系统最大约 qps 为 2 万/秒的并发环境下,全链路最大时延可控制在 30 ms 内。该时延程度对大型应用系统的可用性造成的压力很小,且几乎不会影响到用户体验。如图 7 所示,从线上图数据库查询得到实时动态图平均耗时为 10 ms,最大耗时 15 ms。查询点特征平均耗时 3 ms,最大耗时 5 ms。模型打分平均耗时 5 ms,最大耗时 10 ms。上述 3 项关键步骤累计合计平均耗时 18 ms,最坏情况下耗时 30 ms。

同时,实时动态图框架由于采用了端到端架构,引

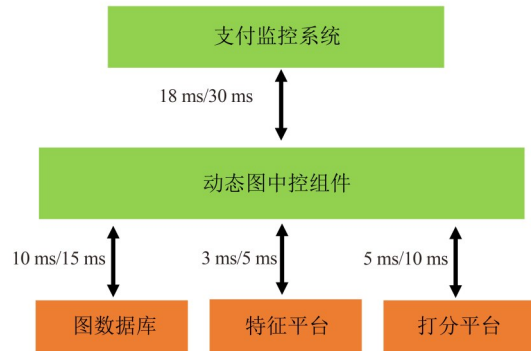


图 7 系统运行时性能

入了大量智能化、自动化学习,成功取代了原框架下的大量天到周级别的人工干预部分。通过对模型智能化的升级,运营性能和成本获得了显著改善。

5.3 实验参数与细节

由于上述实验采用了超大规模场景下的工业数据集,涉及的样本量级和特征维度巨大,在训练过程中采用了基于参数服务器的分布式训练方式。具体来说,使用了 10 个工作节点,2 个参数服务节点。同时,为了对比的公平性,选取了和上述对比基线类似的超参设定。本文采用了常见的 Adam 优化器进行梯度优化,设置学习率为 10^{-4} ,正则化项为 10^{-3} ,批处理大小为 256。为简化起见,对于图中用户和事件的隐层表征,统一将其维度空间设置为 64。值得一提的是,在聚合时使用了 3 层图神经网络架构。通过扩大参与聚合的点集合,增加了模型可获取的信息量并进一步小幅提升了模型效果。对于时间编码函数,选取了和文献[33]相一致的时间编码参数,并将时间的规一化量纲设置为 1 天(86 400 s)。

6 结论与展望

本文提出了实时动态图联合学习框架——RULF,并在实时的金融交易风控场景中取得了良好的效果。为了灵活捕捉图上的高阶结构化时序信息,在时间建模中,使用了函数化的时间编码和连续时间背景感知的图注意力深度网络并进行堆叠。框架设计方面,将金融场景中多角色的联合行为和单角色的独立行为进行了显式的解耦,以使各个模块可以更专注于被赋予的细分目标。最后,将多个模块联合起来学习,并搭建了一套已在工业级成功大规模应用的离在线系统架构。相对于学术模型,本文框架的优势被大量公开可用和真实工业数据集的实验所验证,其有效性也在实际业务场景中得到了证实。在未来的工作中,我们将进一步致力于研究金融场景中数据的稀疏性、非均衡问题,将元学习或多目标学习等技术集成到我们的框架中。

参考文献

- [1] YANG S, ZHANG Z Q, ZHOU J, et al. Financial risk analysis for SMEs with graph-based supply chain mining[C]// Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence. New York: ACM, 2021: 4661-4667.
- [2] CAO S S, YANG X X, CHEN C, et al. TitAnt[J]. Proceedings of the VLDB Endowment, 2019, 12(12): 2082-2093.
- [3] ZHU X Q, AO X, QIN Z D, et al. Intelligent financial fraud detection practices in post-pandemic era[J]. Innovation (Cambridge), 2021, 2(4): 100176.
- [4] ZHU Y, LI H, LIAO Y K, et al. What to do next: Modeling user behaviors by time-LSTM[C]// Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence. Melbourne: IJCAI, 2017: 3602-3608.
- [5] KUMAR S, ZHANG X K, LESKOVEC J. Predicting dynamic embedding trajectory in temporal interaction networks[C]// Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM, 2019: 1269-1278.
- [6] TAHEREH P, KOK-LEONG O, BOOI H KAM, et al. Fraud detection: A systematic literature review of graph-based anomaly detection approaches[J]. Decision Support Systems, 2020, 133: 113303.
- [7] RAWAT W, WANG Z H. Deep convolutional neural networks for image classification: A comprehensive review [J]. Neural Computation, 2017, 29(9): 2352-2449.
- [8] KIPF T N, WELING M. Semi-supervised classification with graph convolutional networks[EB/OL]. (2016)[2022]. <https://arxiv.org/abs/1609.02907.pdf>.
- [9] VELIČKOVIĆ P, CUCURULL G, CASANOVA A, et al. Graph attention networks[EB/OL]. (2017) [2022]. <https://arxiv.org/abs/1710.10903.pdf>.
- [10] LIU Z Q, CHEN C C, LI L F, et al. GeniePath: Graph neural networks with adaptive receptive paths[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33(1): 4424-4431.
- [11] BO D Y, HU B B, WANG X A, et al. Regularizing graph neural networks via consistency-diversity graph augmentations[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2022, 36(4): 3913-3921.
- [12] TU K, CUI P, WANG D X, et al. Conditional graph attention networks for distilling and refining knowledge graphs in recommendation[C]// Proceedings of the 30th ACM International Conference on Information & Knowledge Management. New York: ACM, 2021: 1834-1843.
- [13] ZHANG M M, WANG X A, ZHU M Q, et al. Robust heterogeneous graph neural networks against adversarial attacks[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2022, 36(4): 4363-4370.
- [14] LIU H R, HU B B, WANG X, et al. Confidence may cheat: Self-training on graph neural networks under distribution shift[EB/OL]. (2022)[2022]. <https://arxiv.org/abs/2201.11349.pdf>.
- [15] YANG S, HU B B, ZHANG Z Q, et al. Inductive link prediction with interactive structure learning on attributed graph[C]// Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Cham: Springer, 2021: 383-398.
- [16] GUO S N, LIN Y F, FENG N, et al. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33(1): 922-929.
- [17] HOU Y P, HU B B, ZHAO W X, et al. Neural graph matching for pre-training graph neural networks[EB/OL]. (2022)[2022]. <https://arxiv.org/abs/2203.01597.pdf>.
- [18] LIU Z Q, SHEN Y, CHENG X C, et al. Learning representations of inactive users: A cross domain approach with graph neural networks[C]// Proceedings of the 30th ACM International Conference on Information & Knowledge Management. New York: ACM, 2021: 3278-3282.
- [19] YANG M H, CAO S S, HU B B, et al. IntelliTag: an intelligent cloud customer service system based on tag recommendation[C]// 2021 IEEE 37th International Conference on Data Engineering (ICDE). Piscataway: IEEE, 2021: 2559-2570.
- [20] 高留杰, 赵文, 张君福, 等. G2S: 基于语义块的知识图谱问答语义解析[J]. 电子学报, 2021, 49(6): 1132-1141.
- GAO L J, ZHAO W, ZHANG J F, et al. G2S: Semantic segment based semantic parsing for question answering over knowledge graph[J]. Acta Electronica Sinica, 2021, 49(6): 1132-1141. (in Chinese)
- [21] 车向北, 康文倩, 邓彬, 等. 一种基于图神经网络的SDN路由性能预测模型[J]. 电子学报, 2021, 49(3): 484-491.
- CHE X B, KANG W Q, DENG B, et al. A prediction model of SDN routing performance based on graph neural network[J]. Acta Electronica Sinica, 2021, 49(3): 484-491. (in Chinese)
- [22] ANDERSON T. The Theory and Practice of Online Learning[M]. Athabasca: Athabasca University Press, 2008.
- [23] ZHU Z B, LIU Z Q, JIN G, et al. MixSeq: Connecting

macroscopic time series forecasting with microscopic time series data[EB/OL]. (2021)[2022]. <https://arxiv.org/abs/2110.14354.pdf>.

- [24] QIAN H, WU Q T, ZHANG K, et al. Scope-aware re-ranking with gated attention in feed[C]//Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining. New York: ACM, 2022: 804-812.
- [25] ZHANG K, QIAN H, LIU Q, et al. SIFN: A sentiment-aware interactive fusion network for review-based item recommendation[C]//Proceedings of the 30th ACM International Conference on Information & Knowledge Management. New York: ACM, 2021: 3627-3631.
- [26] DU L, WANG Y, SONG G J, et al. Dynamic network embedding: An extended approach for skip-gram based network embedding[C]//Proceedings of the 27th International Joint Conference on Artificial Intelligence. Stockholm: IJCAI, 2018: 2086-2092.
- [27] SINGER U, GUY I, RADINSKY K. Node embedding over temporal graphs[EB/OL]. (2019)[2022]. <https://arxiv.org/abs/1903.08889.pdf>.
- [28] XU D, RUAN C W, KORPEOGLU E, et al. Inductive representation learning on temporal graphs[EB/OL]. (2020)[2022]. <https://arxiv.org/abs/2002.07962.pdf>.
- [29] HUANG H, FANG Z X, WANG X, et al. Motif-preserving temporal network embedding[C]//Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence. Yokohama: IJCAI-20, 2020: 1237-1243.
- [30] 周俊, 张志强, 曹月恬, 等. 地理位置和时间感知的表示学习框架[J]. 智能系统学报, 2021, 16(5): 909-917.
ZHOU J, ZHANG Z Q, CAO Y T, et al. A geography and time aware representation learning framework[J]. CAAI Transactions on Intelligent Systems, 2021, 16(5): 909-917. (in Chinese)
- [31] ZHANG D L, HUANG X, LIU Z Q, et al. AGL: A scalable system for industrial-purpose graph machine learning [J]. Proceedings of the VLDB Endowment. 2020, 13(12): 3125-3137.
- [32] HAMILTON W L, YING R, LESKOVEC J. Inductive representation learning on large graphs[EB/OL]. (2017) [2022]. <https://arxiv.org/abs/1706.02216.pdf>.
- [33] NGUYEN G H, LEE J B, ROSSI R A, et al. Continuous-time dynamic network embeddings[C]//Proceedings of the The Web Conference 2018. New York: ACM, 2018: 969-976.
- [34] LU Y F, WANG X, SHI C, et al. Temporal network em-

bedding with micro- and macro-dynamics[C]//Proceedings of the 28th ACM International Conference on Information and Knowledge Management. New York: ACM, 2019: 469-478.

- [35] SEO Y, DEFFERRARD M, VANDERGHEYNST P, et al. Structured sequence modeling with graph convolutional recurrent networks[C]//CHENG L, LEUNG A, OZAWA S. International Conference on Neural Information Processing. Cham: Springer, 2018: 362-373.

作者简介



周俊男, 1986年出生, 湖南人. 蚂蚁科技集团机器智能部负责人. 主要研究方向为机器学习、图神经网络等. 中国电子学会会员编号: E190028162S.

E-mail: jun.zhoujun@antfin.com



陈超超(通讯作者) 男, 1988年出生, 河南人. 浙江大学特聘研究员. 主要研究方向为图神经网络、推荐系统等.

E-mail: zjuccc@zju.edu.cn