

借这篇博客记录图学习的笔记。

摘要

Temporal Graph Learning

Temporal Heterogeneous Information Network Embedding[IJCAI'21]
DySAT: Deep Neural Representation Learning on Dynamic Graphs via Self-Attention Networks[WSDM'20]
EvolveGCN Evolving Graph Convolutional Networks for Dynamic Graphs[AAAI'20]
Temporal-Aware Graph Neural Network for Credit Risk Prediction[SDM'21]
T-GCN: A Temporal Graph Convolutional Network for Traffic Prediction[TITS'17]
Characterizing and Forecasting User Engagement with In-app Action Graph: A Case Study of Snapchat[KDD'19]
Modeling Dynamic Heterogeneous Network for Link Prediction using Hierarchical Attention with Temporal RNN[ECML-PKDD'20]

Heterogeneous Graph Learning

Heterogeneous Graph Attention Network[WWW'19]
Are we really making much progress? Revisiting, benchmarking, and refining heterogeneous graph neural networks[KDD'21]

Graph Anomaly Detection

Deep Anomaly Detection on Attributed Networks[SDM'19]
GeniePath: Graph Neural Networks with Adaptive Receptive Paths[AAAI'19]
Heterogeneous Graph Neural Networks for Malicious Account Detection[CIKM'18]
Decoupling Representation Learning and Classification for GNN-based Anomaly Detection[SIGIR'21]
AddGraph: Anomaly Detection in Dynamic Graph Using Attention-based Temporal GCN[IJCAI'19]
Structural Temporal Graph Neural Networks for Anomaly Detection in Dynamic Graphs[CIKM'20]
Anomaly Detection in Dynamic Graphs via Transformer[TKDE'21]

Graph Neural Networks

How Attentive are Graph Attention Networks?[ICLR'22]

Sequential Recommendation with Contrastive Learning

Contrastive Learning for Sequential Recommendation[SIGIR'21]
Contrastive Learning for Representation Degeneration Problem in Sequential Recommendation[WSDM'22]
Improving Sequential Recommendation Consistency with Self-Supervised Imitation[IJCAI'21]
Intent Contrastive Learning for Sequential Recommendation[WWW'22]
Sequential Recommendation with Multiple Contrast Signals[TOIS'22]
Contrastive Self-supervised Sequential Recommendation with Robust Augmentation[arxiv'21]
Augmenting Sequential Recommendation with Pseudo-Prior Items via Reversely Pre-training Transformer[SIGIR'21]
Sequential Recommendation with Bidirectional Chronological Augmentation of Transformer self-supervised learning for sequential recommendation with model augmentation

Graph Self-Supervised Learning

Self-supervised Graph Learning for Recommendation[SIGIR'21]
Self-supervised Representation Learning on Dynamic Graphs[CIKM'21]
Multi-View Self-Supervised Heterogeneous Graph Embedding[ECML-PKDD'21]
Graph Debaised Contrastive Learning with Joint Representation Clustering[IJCAI'21]
GCC--Graph Contrastive Coding for Graph Neural Network Pre-Training[KDD'20]
GCCAD: Graph Contrastive Coding for Anomaly Detection
Deep Graph Infomax[ICLR'19]
InfoGraph: Unsupervised and Semi-supervised Graph-Level Representation Learning via Mutual Information Maximization[ICLR'20]
Spatio-Temporal Deep Graph Infomax[ICLR'19 workshop]
Self-Supervised Hypergraph Convolutional Networks for Session-based Recommendation[AAAI'21]
Self-Supervised Multi-Channel Hypergraph Convolutional Network for Social Recommendation[WWW'21]

Contrastive and Generative Graph Convolutional Networks for Graph-based Semi-Supervised Learning[AAAI'21]
Spatio-Temporal Graph Contrastive Learning
TCL: Transformer-based Dynamic Graph Modelling via Contrastive Learning
Structure-Enhanced Heterogeneous Graph Contrastive Learning
Contrastive Pre-Training of GNNs on Heterogeneous Graphs[CIKM'21]
Self-supervised Heterogeneous Graph Neural Network with Co-contrastive Learning[KDD'21]

摘要

Temporal Graph Learning

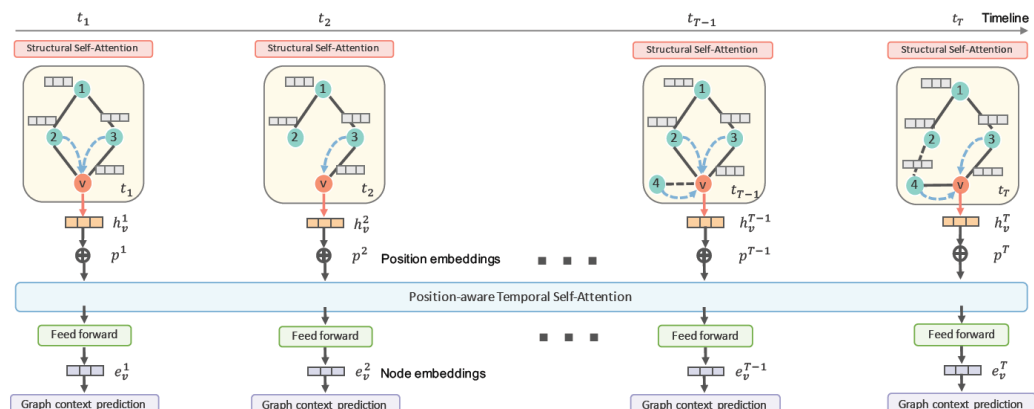
动态图学习相关论文，对每篇读过的论文给出基本介绍

Temporal Heterogeneous Information Network Embedding[IJCAI'21]

- 标签：动态图学习、异构图
- 摘要：对于异构信息（顶点和边），论文使用基于元路径的随机游走来处理。元路径能够描述起始顶点和终止顶点之间的复杂关系，它可以经过任意类型的顶点和边。对于动态信息，使用Hawkes过程来建模过去事件对当前事件的影响，以attention机制来区分每个事件影响的重要程度。
- 链接：<https://www.ijcai.org/proceedings/2021/203>
- 是否有开源代码：<https://github.com/S-rz/THINE>

DySAT: Deep Neural Representation Learning on Dynamic Graphs via Self-Attention Networks[WSDM'20]

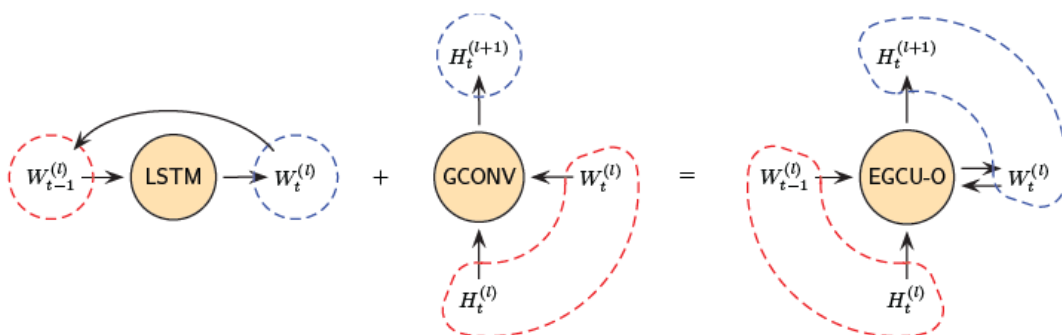
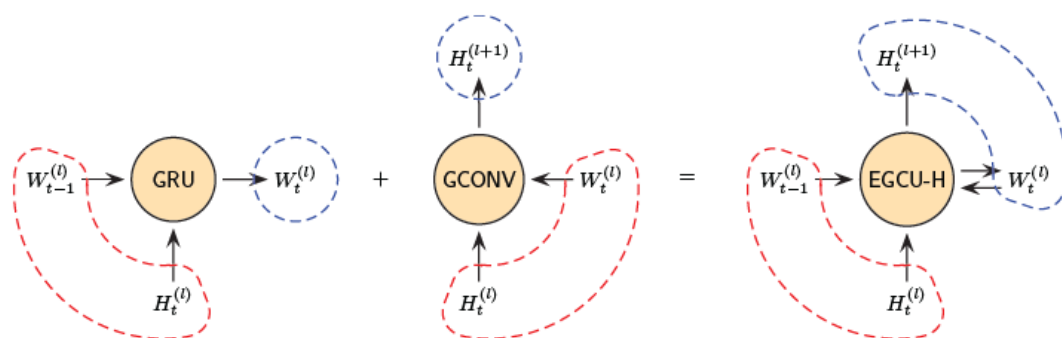
- 标签：attention
- 摘要：论文提出DySAT模型，从结构和时序两个方面应用注意力机制。首先在各个快照图上，通过结构上的注意力机制加权的形式对邻域信息进行聚合，得到当前时刻下图上各顶点的特征表示；随后将顶点在不同时刻下的表示，加上position embedding，拼接成一个序列，通过时序上的注意力机制，将输入序列转换为一个新的输出序列。序列中各快照之间的注意力系数定义为，



- 链接：<https://dl.acm.org/doi/10.1145/3336191.3371845>
- 是否有开源代码：https://github.com/FeiGSSS/DySAT_pytorch

EvolveGCN Evolving Graph Convolutional Networks for Dynamic Graphs[AAAI'20]

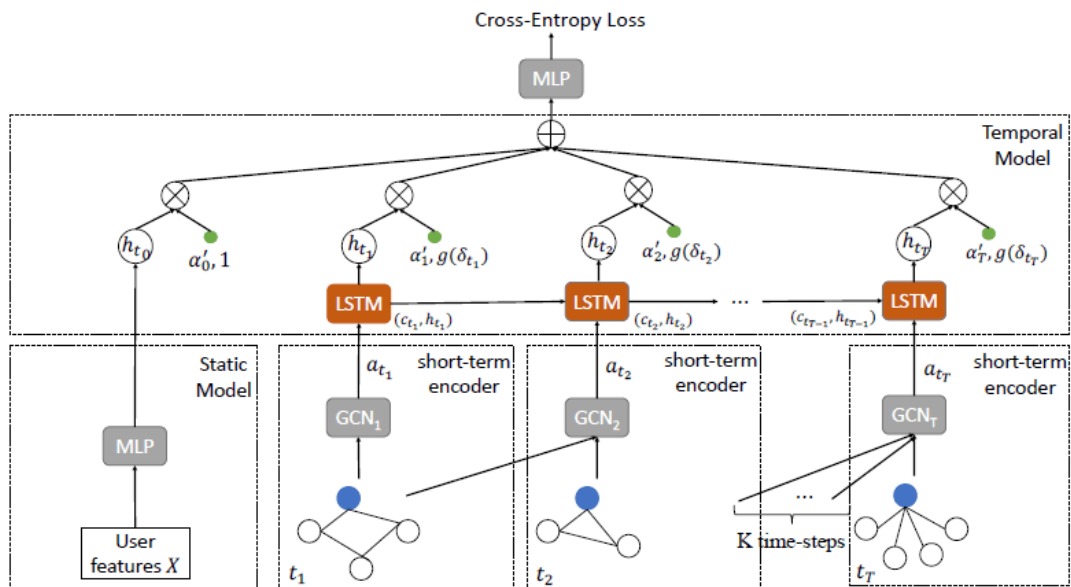
- 标签：动态图学习
- 摘要：将动态图看做一系列快照图，EvolveGCN将 $t+1$ 时刻与 t 时刻特征之间的关系通过GCN表示为 $H_t^{(l+1)} = \sigma(A_t H_t^{(l)} W_t^{(l)})$ ，关键在于如何更新 $W_t^{(l)}$ 。第一种方式将它看作网络的隐层表示，通过GRU进行更新；第二种方式将它看作网络的输出，通过LSTM进行更新。前者额外的将 $H_T^{(l)}$ 作为输入，而后者将上一时刻的输出直接作为输入：



- 链接：<https://ojs.aaai.org/index.php/AAAI/article/view/5984/5840>
- 是否有开源代码：<https://github.com/IBM/EvolveGCN>

Temporal-Aware Graph Neural Network for Credit Risk Prediction[SDM'21]

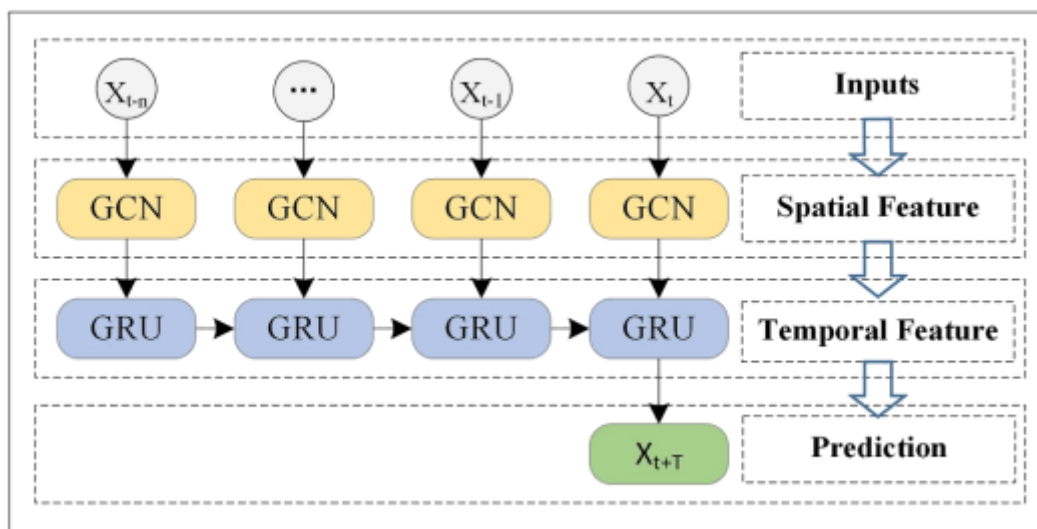
- 标签：动态图学习
- 摘要：蚂蚁金服提出TemGNN模型用于信用卡欺诈检测。模型分为三部分，第一部分为MLP，用于将固定不变的特征如教育背景转换为静态的特征向量；第二部分为short-term图卷积，通过滑动窗口+attention聚合的形式对邻域信息进行聚合；第三部分为long-term编码，相较上一部分使用一个更长的滑动窗口，将形成的快照图序列通过LSTM，但TemGNN没有直接将LSTM的最终输出作为结果，而是将每一个cell的输出通过一个attention的形式进行聚合，其中还加入了一个时间衰减因子，来减少久远事件对当前的影响。



- 链接: <https://epubs.siam.org/doi/abs/10.1137/1.9781611976700.79?mobileUi=0>
- 是否有开源代码: 无

T-GCN: A Temporal Graph Convolutional Network for Traffic Prediction[TITS'17]

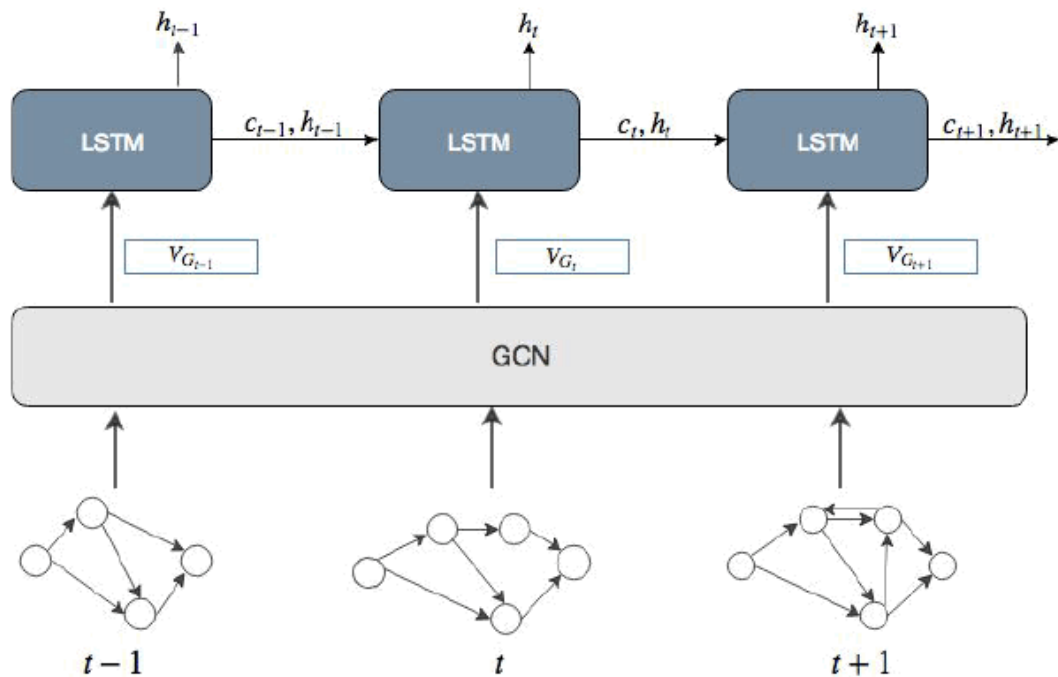
- 标签:
- 摘要: T-GCN将GCN与GRU结合来捕获交通数据中的空间与时间信息, 其中GCN用于学习空间上的拓扑结构而GRU用来学习时间上的动态变化。



- 链接: <https://ieeexplore.ieee.org/document/8809901>
- 是否有开源代码: <https://github.com/lehaifeng/T-GCN/tree/master/T-GCN/T-GCN-PyTorch>

Characterizing and Forecasting User Engagement with In-app Action Graph: A Case Study of Snapchat[KDD'19]

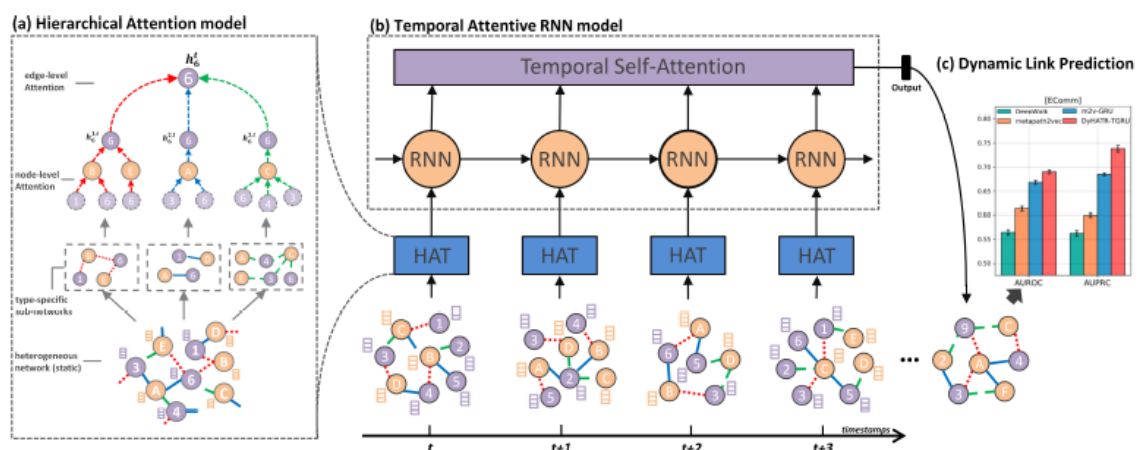
- 标签: 动态图学习
- 摘要: 在建模动态信息时, 论文将GCN与LSTM相结合, 先通过GCN为每个时刻的快照图生成 embedding, 所有时刻的快照图embedding就形成了一条序列, 再将这条序列输入LSTM中, 取最后一层的隐层输出为最终的特征表示。



- 链接: <https://dl.acm.org/doi/10.1145/3292500.3330750>
- 是否有开源代码: <https://github.com/INK-USC/temporal-gcn-lstm>

Modeling Dynamic Heterogeneous Network for Link Prediction using Hierarchical Attention with Temporal RNN[ECML-PKDD'20]

- 标签:
- 摘要: 论文提出的DyHATR是一个动态异构图的框架。在每个快照图上，通过两个层面的attention得到顶点的特征表示：顶点层面attention在各异构边形成的子图中，通过attention聚合邻域信息，边层面attention再将各异构边聚合的信息通过attention加权聚合得到结果。将顶点在各时刻下的特征表示合起来看作一条序列，通过RNN+self-attention的形式在时序上使用attention加权聚合各时刻的信息，得到最终的顶点表示用于下游任务。



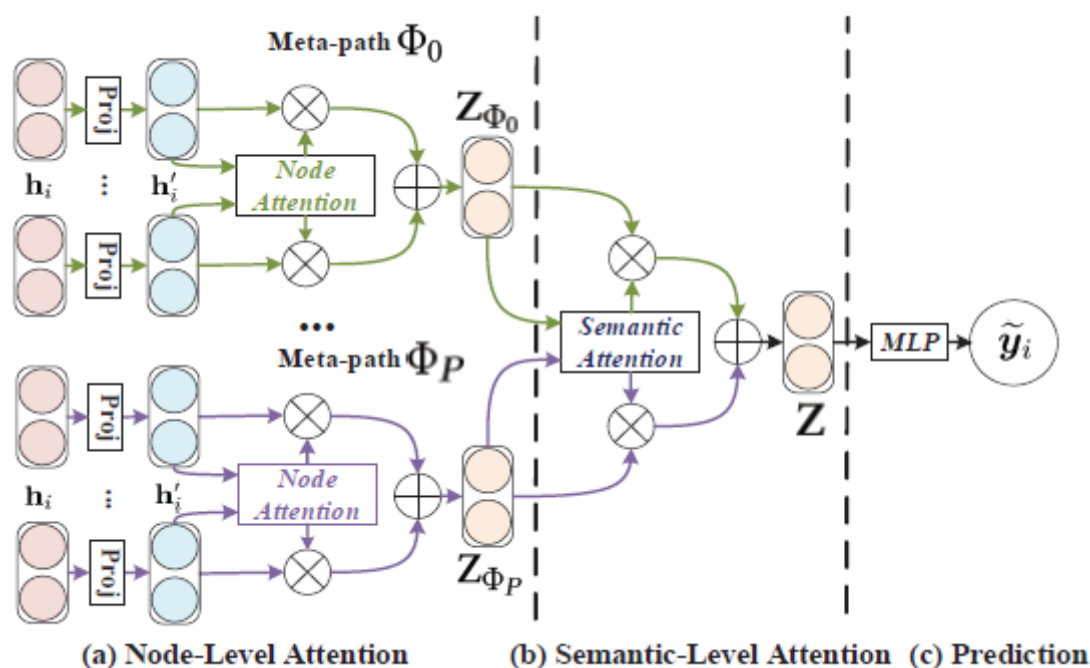
- 链接: https://link.springer.com/chapter/10.1007/978-3-030-67658-2_17
- 是否有开源代码: <https://github.com/skx300/DyHATR>

Heterogeneous Graph Learning

异构图学习相关论文，对每篇读过的论文给出基本介绍

Heterogeneous Graph Attention Network[WWW'19]

- 标签：异构图
- 摘要：论文提出HAN模型，使用了两个层次的attention：node-level和semantic-level。node-level是指在特定的元路径下聚合邻域信息时学习重要性，semantic-level是指学习不同元路径的重要性。



- 链接：<https://dl.acm.org/doi/10.1145/3308558.3313562>
- 是否有开源代码：<https://github.com/dmlc/dgl/tree/master/examples/pytorch/han>

Are we really making much progress? Revisiting, benchmarking, and refining heterogeneous graph neural networks[KDD'21]

- 标签：异构图
- 摘要：论文在GAT的基础上提出了异构图上的模型Simple-HGN，加入了三点改进：边 embedding、顶点和边的残差连接以及 L_2 正则化：
 - 在计算attention系数时，GAT根据顶点和邻域的特征表示计算，Simple-HGN进一步加入了边类型的embedding：

$$\hat{\alpha}_{ij} = \frac{\exp \left(\text{LeakyReLU} \left(a^T [W h_i \| W h_j \| W_r r_{\psi(\langle i, j \rangle)}] \right) \right)}{\sum_{k \in \mathcal{N}_i} \exp \left(\text{LeakyReLU} \left(a^T [W h_i \| W h_k \| W_r r_{\psi(\langle i, k \rangle)}] \right) \right)}$$

- 顶点残差连接指的是，在得到顶点 i 在第 l 层的特征表示时，除了聚合邻域顶点在上一层的特征表示之外，还额外加上上一层的输出：

$$h_i^{(l)} = \sigma \left(\sum_{j \in \mathcal{N}_i} \alpha_{ij}^{(l)} W^{(l)} h_j^{(l-1)} + h_i^{(l-1)} \right)$$

除此之外，每一层的attention系数也会与上一层做加权平均，构成所谓的边残差连接。

◦ L_2 正则就是字面意义上的在输出上进行 L_2 正则化。

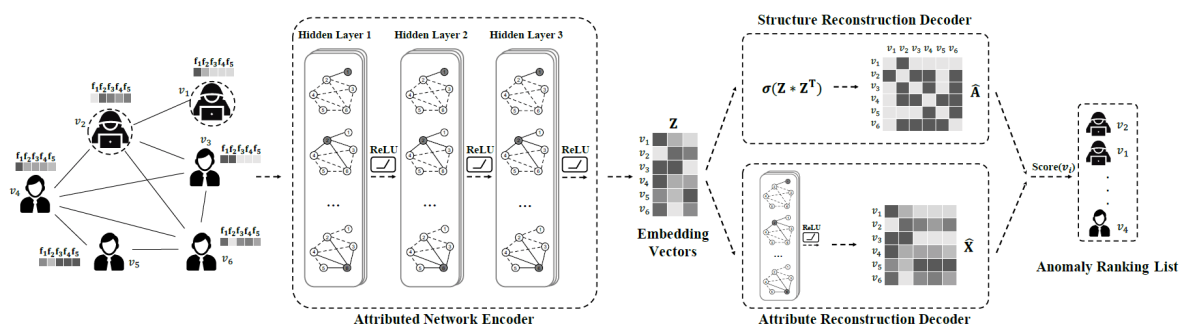
- 链接: <https://dl.acm.org/doi/10.1145/3447548.3467350>
- 是否有开源代码: <https://github.com/BUPT-GAMMA/OpenHGNN/blob/main/openhgnn/models/SimpleHGN.py>

Graph Anomaly Detection

图异常检测相关论文，对每篇读过的论文给出基本介绍

Deep Anomaly Detection on Attributed Networks[SDM'19]

- 标签: 矩阵分解
- 摘要: 论文提出的DOMINANT方法十分直观，是一个典型的Encoder-Decoder架构。首先通过GCN对输入图进行编码，再将得到的顶点embedding通过向量点乘的形式去还原输入图的邻接矩阵与特征矩阵，通过近似值与真实值之间的差值计算损失函数，以及计算目标顶点的预测分数。

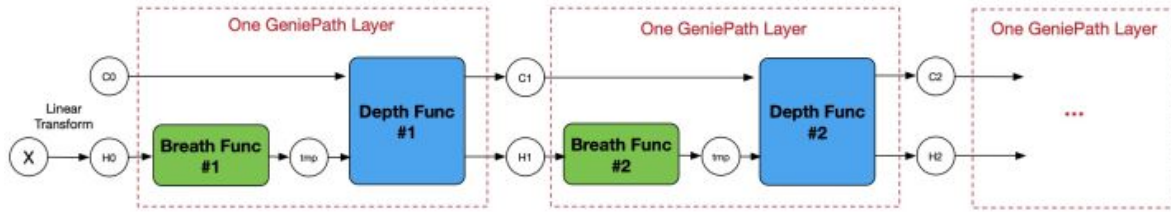


- 链接: <https://epubs.siam.org/doi/abs/10.1137/1.9781611975673.67?mobileUi=0>
- 是否有开源代码: https://github.com/kaize0409/GCN_AnomalyDetection_pytorch

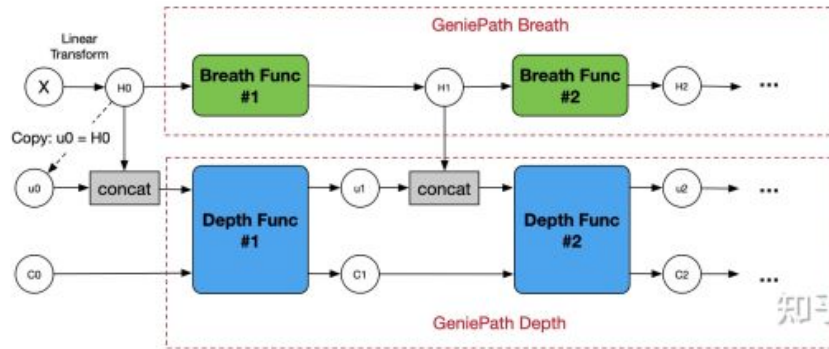
GeniePath: Graph Neural Networks with Adaptive Receptive Paths[AAAI'19]

- 标签:
- 摘要: 参考 GeniePath 是蚂蚁团队提出的一个在图上自适应地聚合邻域信息的方法，每一层都由 breadth 和 depth 两部分组成，前者通过 GAT 为一阶邻居赋予不同的重要性，后者基于 LSTM，将 t 阶邻居看作 t 个时刻，通过遗忘门和保留门来对高阶邻居的信息进行过滤，以达到自适应的目的。

GeniePath-Standard



GeniePath-Lazy



- 链接: <https://aaai.org/ojs/index.php/AAAI/article/view/4354/4232>
- 有无开源代码: <https://github.com/kayzliu/GeniePath>

Heterogeneous Graph Neural Networks for Malicious Account Detection[CIKM'18]

- 标签: 异构图
- 摘要: 论文提出GEM模型, 通过观察恶意账户的行为特征来构建异构图。顶点类型有用户和设备(ip地址、手机号码等), 边类型为用户对设备的行为(注册、登录等)。GEM对输入图从不同设备的角度提取D个子图, 在构建顶点的特征时, 特征向量的维度为 $p+D$, 前p维是用户顶点特征, 将用户在 $0 \sim T$ 时间内的行为划分为p个时间段, 每个时间段有一个行为次数, 而后D维是设备顶点特征, 是所属子图的one-hot编码。提出的GEM模型将各个子图作为邻域进行聚合:

$$H^{(0)} \leftarrow 0$$

$$\text{for } t = 1, \dots, T$$

$$H^{(t)} \leftarrow \sigma \left(X \cdot W + \frac{1}{|\mathcal{D}|} \sum_{d=1}^{|\mathcal{D}|} A^{(d)} \cdot H^{(t-1)} \cdot V_d \right)$$

额外地, GEM在聚合时使用attention来学习各子图的重要性。

- 链接: <https://dl.acm.org/doi/10.1145/3269206.3272010>
- 是否有开源代码: <https://github.com/safe-graph/DGFraud-TF2/tree/main/algorithms/GEM>

Decoupling Representation Learning and Classification for GNN-based Anomaly Detection[SIGIR'21]

- 标签：聚类
- 摘要：论文提出了DCI方法，在DGI基础上加入了k-means聚类。DGI通过顶点-图的形式进行对比学习，来训练编码器，其中图的表示通过pooling得到。而DCI通过顶点-子图的形式进行对比学习，其中的子图由k-means得到，总的损失函数由各子图取平均得到。

Algorithm 1: Deep Cluster Infomax

Input : Graph $G = (V, A, X)$, Number of clusters K ,
Number of training epochs t , Number of
re-clustering epochs \bar{t} .

Output: Optimized GNN encoder g

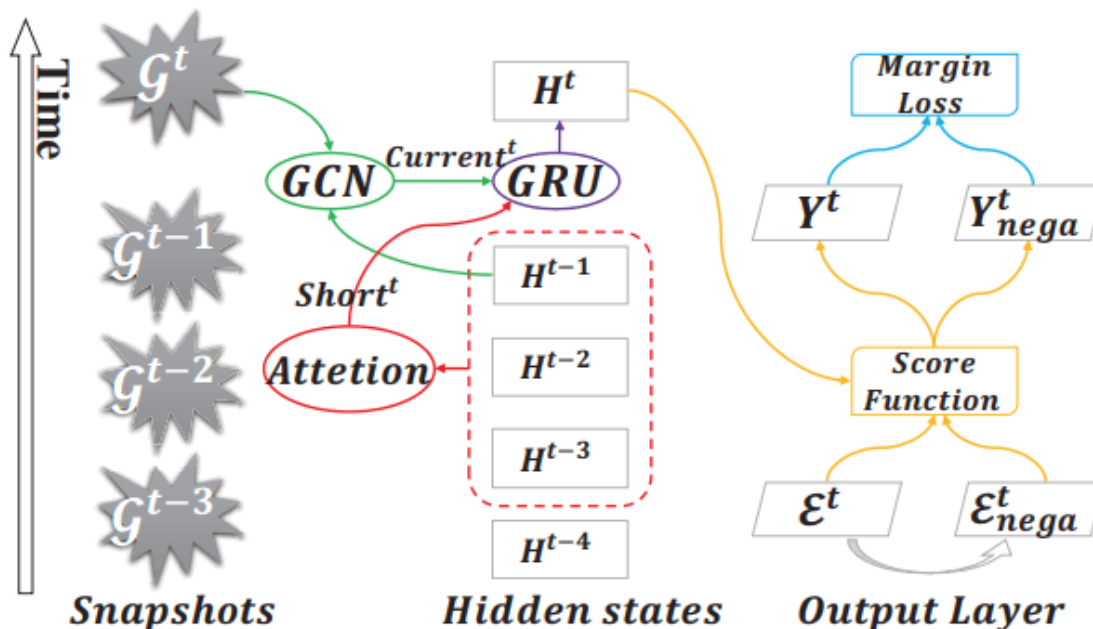
```
1 Initialize clusters  $[C_1, C_2, \dots, C_K] = \text{K-Means}(X)$ ;  
2 Initialize the parameters  $\theta$  and  $\omega$  for the encoder  $g$  and  
   the discriminator  $\mathcal{D}$  ;  
3 for  $epoch \leftarrow 1$  to  $t$  do  
4    $H = g(G, \theta)$ ;  
5    $\mathcal{L}_{DCI} = \frac{1}{K} \sum_{k=1}^K \mathcal{L}_{DCI}^k(H, C_k, \omega)$ ;  
6    $\theta, \omega \leftarrow \text{Adam}(\mathcal{L}_{DCI})$ ;  
7   if  $t \bmod \bar{t} == 0$  then  
8      $[C_1, C_2, \dots, C_K] = \text{K-Means}(g(G, \theta))$ 
```

Return: encoder g

- 链接: <https://dl.acm.org/doi/10.1145/3404835.3462944>
- 是否有开源代码: <https://github.com/wyl7/DCI-pytorch>

AddGraph: Anomaly Detection in Dynamic Graph Using Attention-based Temporal GCN[IJCAI'19]

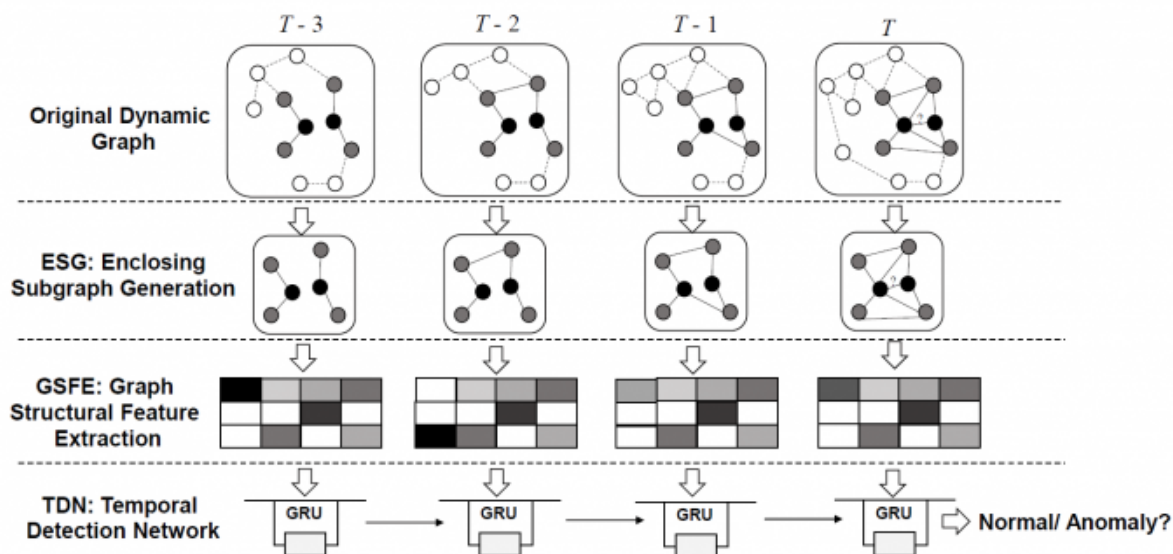
- 标签：动态图学习
- 摘要：论文的目的是检测图上的异常边。做法类似于T-GCN，对于一个包含若干快照的动态图，为了得到顶点在 t 时刻的特征表示，AddGraph首先将 $t-1$ 时刻的特征表示通过一个多层的GCN，得到长期的表示，再将 t 时刻的顶点通过一个基于attention的上下文模型来得到短期的表示，将长短期两部分输入一个GRU来得到 t 时刻的特征表示 H^t ，随后被用来计算两个顶点间出现异常边的概率。因为有标签数据的获取难度，论文还加入了SVM中的Hinge损失函数来进行负采样。



- 链接: <https://www.ijcai.org/proceedings/2019/614>
- 是否有开源代码: <https://github.com/Ljiajie/Addgraph>

Structural Temporal Graph Neural Networks for Anomaly Detection in Dynamic Graphs[CIKM'20]

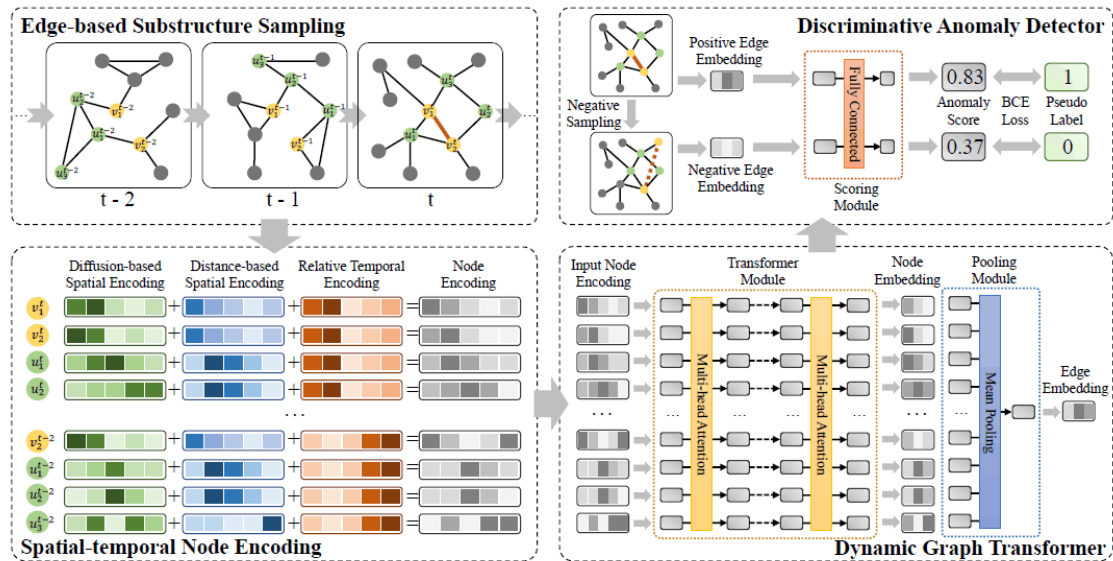
- 标签: 动态图学习
- 摘要: 论文提出的StrGNN用于异常边检测, 分为三个部分: 第一部分以边为中心提取h-hop子图, 并根据结构信息对子图中的顶点进行标记; 第二部分通过GNN为上一步中的子图编码, 并通过 Sort pooling的方式只取子图中的重要性前k个顶点形成子图的特征表示; 第三部分通过GRU将各时刻特征表示形成的序列合成为t时刻目标边的特征表示。



- 链接: <https://dl.acm.org/doi/pdf/10.1145/3459637.3481955>
- 是否有开源代码: 无

Anomaly Detection in Dynamic Graphs via Transformer[TKDE'21]

- 标签: 动态图学习
- 摘要: 论文提出了TADDY方法, 将Transformer应用于动态图的异常边检测问题。TADDY分为四个部分, 第一部分为以边为中心提取子图, 不同于StrGNN采用h-hop的形式, TADDY通过graph diffusion如PPR的形式为图中顶点计算connectivity, 取目标边的头尾顶点中connectivity值top-k的顶点来构造子图。第二部分为子图中的顶点进行编码: diffusion encoding取connectivity值排序后的顺序来编码; distance encoding取顶点到目标边头尾顶点的最短距离来编码; temporal encoding取目标边时间戳与当前快照时间戳差值来编码。第三部分为动态Transformer, 将上一步的编码作为特征输入, 输出的子图中顶点embedding通过pooling得到目标边的embedding。最后一部分将边的embedding转换为对应的异常分数。



- 链接: <https://shiruihan.github.io/publication/tkde-liu-21/>
- 是否有开源代码: https://github.com/yuetan031/TADDY_pytorch

Graph Neural Networks

图神经网络相关论文, 对每篇读过的论文给出基本介绍

How Attentive are Graph Attention Networks?[ICLR'22]

- 标签: attention
- 摘要: 论文提出GATv2模型, 作为GAT的改进, 改动的地方十分简单直观, 将原先GAT中组合在一起的 $a^T W$ 中的移动到非线性层LeakyReLU之后:

$$\text{GAT (Velićković et al., 2018): } e(h_i, h_j) = \text{LeakyReLU}(a^T \cdot [W h_i \| W h_j])$$

$$\text{GATv2 (our fixed version): } e(h_i, h_j) = a^T \text{LeakyReLU}(W \cdot [h_i \| h_j])$$

论文的主要内容就是论述这种改动是如何将GAT的静态attention变成动态attention的。

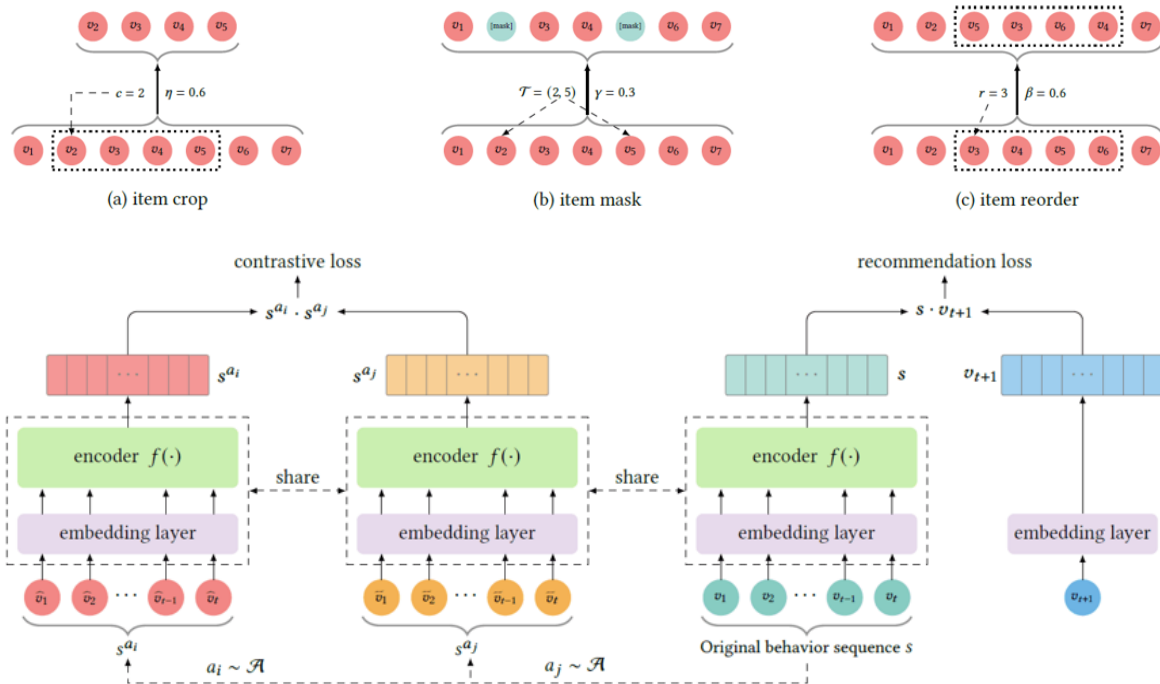
- 链接: <https://openreview.net/forum?id=F72ximsx7C1>
- 是否有开源代码: https://github.com/tech-srl/how_attentive_are_gats

Sequential Recommendation with Contrastive Learning

对比学习在序列推荐任务中的应用, 对每篇读过的论文给出基本介绍

Contrastive Learning for Sequential Recommendation[SIGIR'21]

- 标签：序列推荐
- 摘要：论文提出了三种序列的数据增强策略，分别为：裁剪：取原序列的一段连续子序列；遮盖：随机遮盖序列中部分元素；乱序：将原序列随机乱序，对同一序列进行两次数据增强，得到两条增强序列，在一个batch中，同一序列的两条增强序列为正样本，剩余来自不同序列的增强序列为负样本。



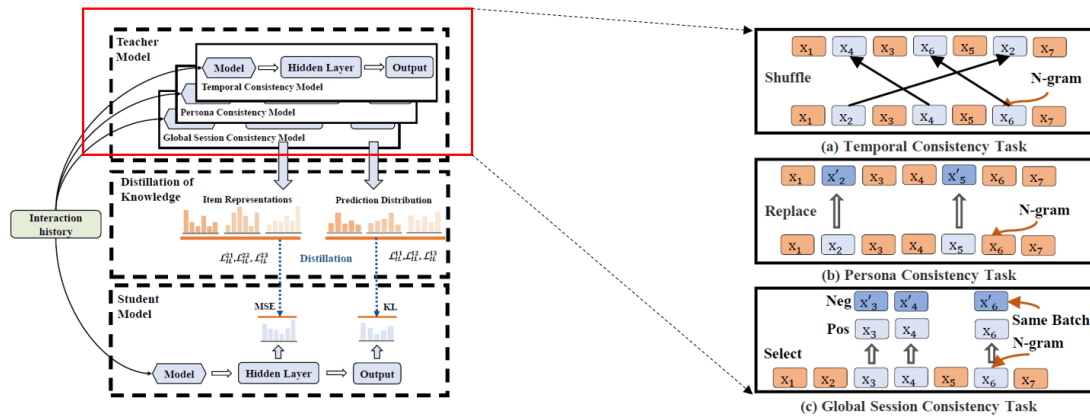
- 链接：<https://arxiv.org/abs/2010.14395>
- 是否有开源代码：<https://github.com/RuihongQiu/DuoRec>

Contrastive Learning for Representation Degeneration Problem in Sequential Recommendation[WSDM'22]

- 标签：序列推荐
- 摘要：无监督增强：论文从模型层面进行数据增强，embedding层与encoder层通过不同的dropout产生增强样本；有监督采样：具有相同推荐商品的序列作为正样本，同一Batch内其余序列作为负样本
- 链接：<https://arxiv.org/abs/2110.05730>
- 是否有开源代码：<https://github.com/RuihongQiu/DuoRec>

Improving Sequential Recommendation Consistency with Self-Supervised Imitation[IJCAI'21]

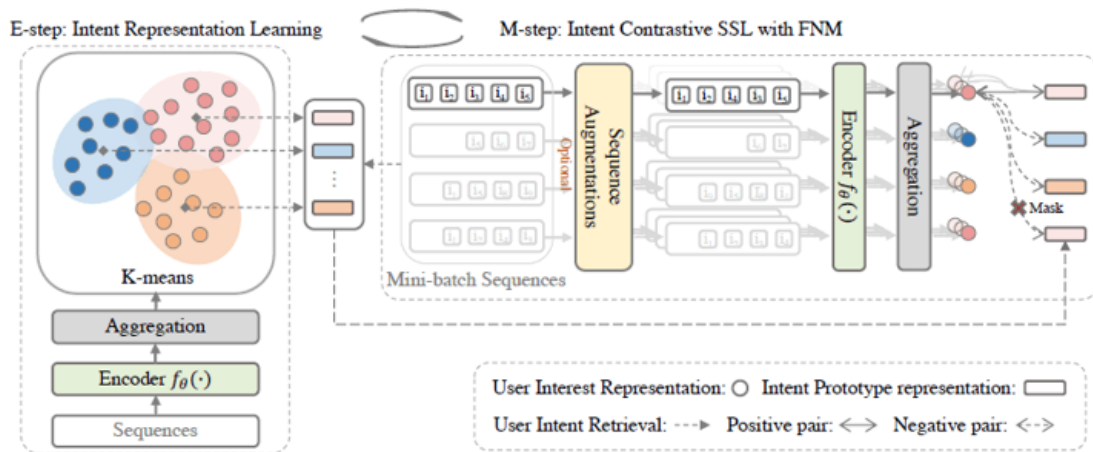
- 标签：预训练
- 摘要：采用教师-学生模型的知识蒸馏方式，对比学习体现在教师模型的两个预训练任务：时序性：将原序列随机乱序；用户兴趣一致性：将序列部分商品随机替换成其他商品；全局一致性：取子序列与整个序列做局部-全局对比。



- 链接: <https://www.ijcai.org/proceedings/2021/0457.pdf>
- 是否有开源代码: 无

Intent Contrastive Learning for Sequential Recommendation[WWW'22]

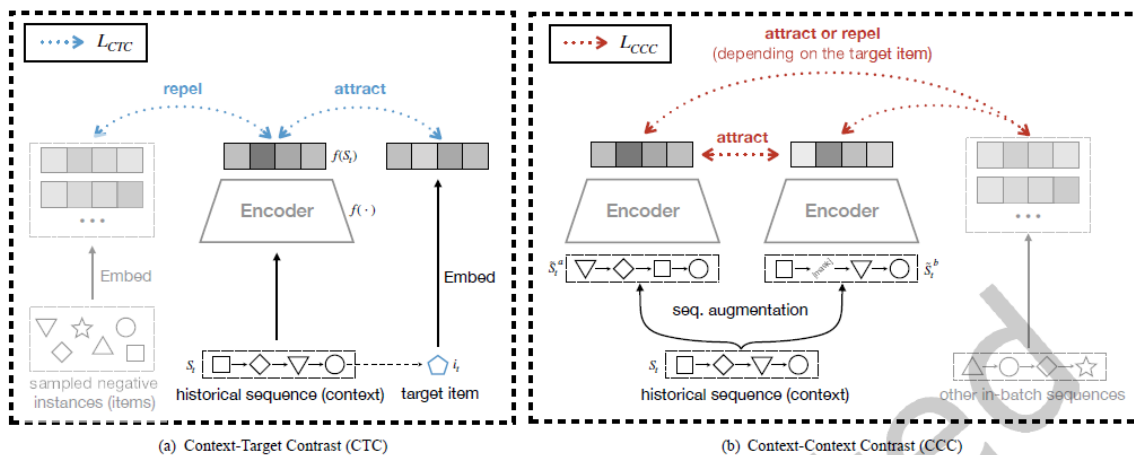
- 标签: 序列推荐
- 摘要: 将序列与其中的隐变量-用户偏好做对比学习, 序列增强是可选项, 即可以通过原序列或增强后的序列来做对比学习, 用户偏好的表示通过聚类形成的簇中心得到, 同一簇内为正样本, 不同簇内为负样本。



- 链接: <https://arxiv.org/abs/2202.02519>
- 是否有开源代码: <https://github.com/YChen1993/ICLRec>

Sequential Recommendation with Multiple Contrast Signals[TOIS'22]

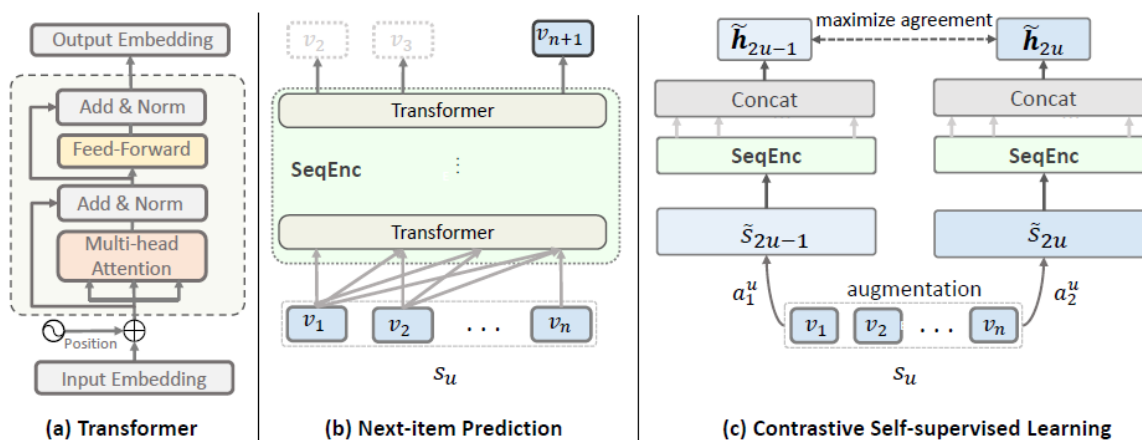
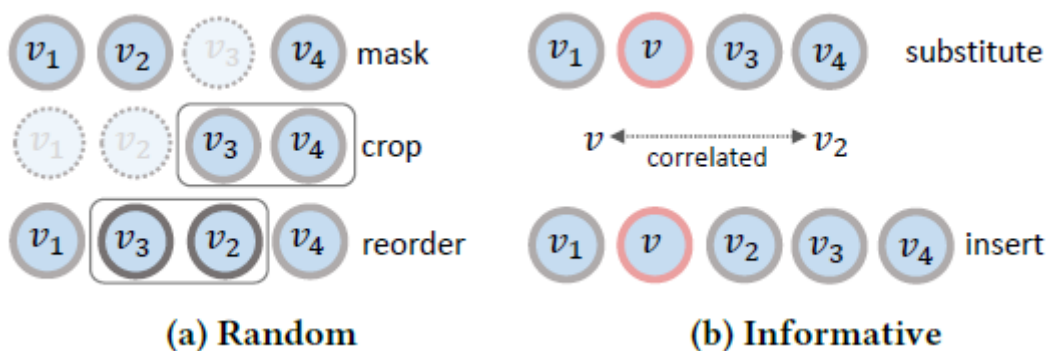
- 标签:
- 摘要: 论文提出的ContraRec由两部分对比任务组成。传统的序列推荐任务通过输入的历史交互序列建模用户的兴趣偏好, 从而预测下一时刻用户可能喜欢的物品, 推荐物品的排序要比其他候选物品更高。论文将这一任务也看作是对比学习的一种形式: 历史序列与推荐物品为正样本, 与其他候选物品为负样本, BPR损失函数也可以很容易地转化为InfoNCE损失函数。除此之外, 传统的序列对比学习作为一个另一个对比任务, 一条序列通过乱序和随机mask两种方式得到两条增强序列, 同一条序列的增强序列自然互为正样本, 而不同序列如果推荐的目标物品相同, 也看做正样本, 否则为负样本。



- 链接: <https://dl.acm.org/doi/abs/10.1145/3522673>
- 是否有开源代码: <https://github.com/THUwangcy/ReChorus/tree/TOIS22>

Contrastive Self-supervised Sequential Recommendation with Robust Augmentation[arxiv'21]

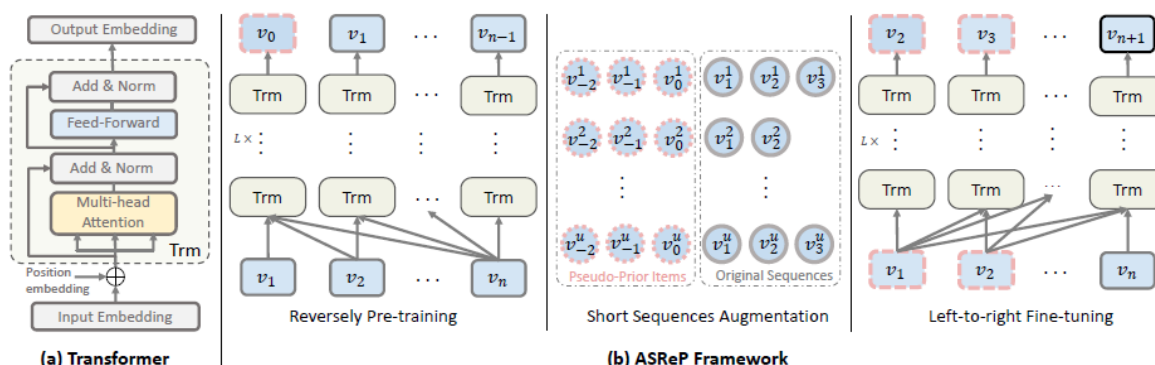
- 标签:
- 摘要: 论文认为传统的序列对比学习中的增强策略如reorder、mask和crop对短序列不适用, 它们破坏了序列内元素的相关性, 或者减少了序列的元素个数加剧了冷启动的问题。本文的主要贡献是提出了两种新的增强策略: 插入insert和替换substitute, 前者是将相关物品j插入序列中物品i的前面, 后者直接用j替换i。衡量物品相关性采用了基于物品的协同过滤和特征向量点乘, 取两种方法归一化后的最大值为最终结果。



- 链接: <https://arxiv.org/abs/2108.06479>
- 是否有开源代码: <https://github.com/YChen1993/CoSeRec>

Augmenting Sequential Recommendation with Pseudo-Prior Items via Reversely Pre-training Transformer[SIGIR'21]

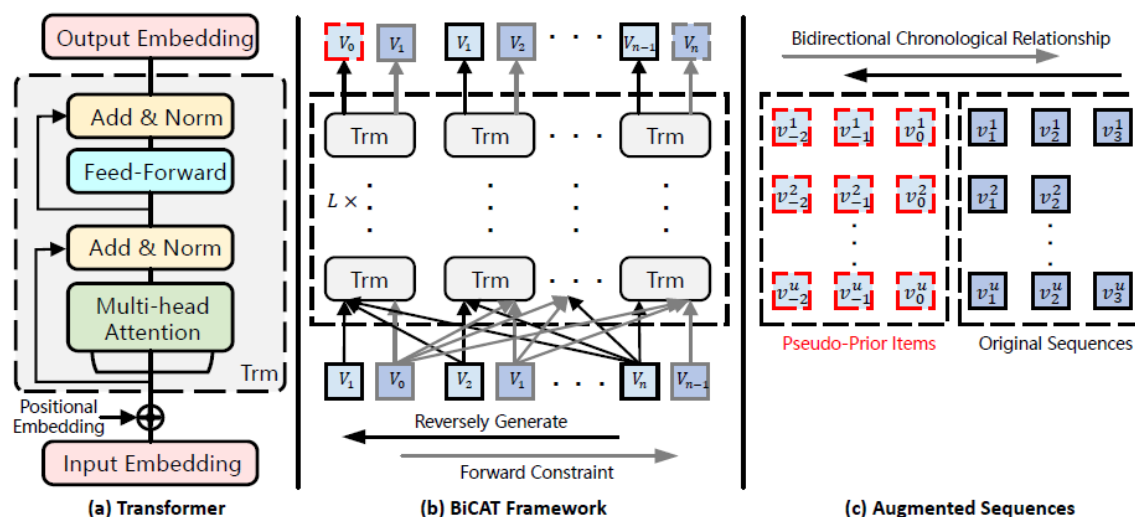
- 标签:
- 摘要: 论文专门研究短序列上的序列推荐任务及对比学习的应用, 提出了ASReP方法。首先将Transformer在输入的序列上逆序进行预训练, 预训练后的模型将生成若干伪商品放置于短序列头部进行拼接, 从而增加序列长度。最后再正序输入Transformer中进行传统的目标商品预测任务。



- 链接: <https://dl.acm.org/doi/abs/10.1145/3404835.3463036>
- 是否有开源代码: <https://github.com/DyGRec/ASReP>

Sequential Recommendation with Bidirectional Chronological Augmentation of Transformer

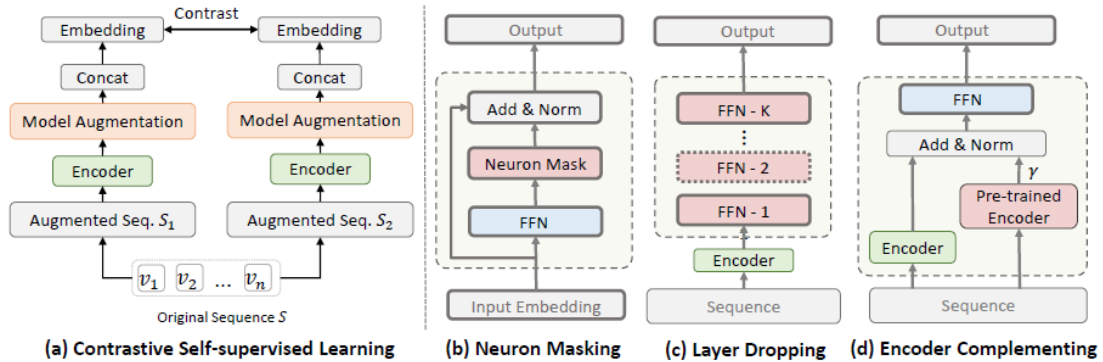
- 标签:
- 摘要: 论文在ASReP的基础上进行改进, 改动的是输入序列逆序来预训练Transformer的部分。预训练任务是传统的目标商品预测任务, 只是逆序情况下预测的目标商品为原序列中第一个商品前面的商品。原序列中商品的下标为 $[1, n]$, 新生成的序列中商品的下标为 $[0, n]$, 它会继续输入相同的Transformer中, 取下标 $[0, n-1]$ 为历史序列, n 为目标商品来进行正序的预训练。在对Transformer预训练后, 后续步骤和ASReP一致, 对短序列增强并预测目标商品。



- 链接: <https://arxiv.org/abs/2112.06460v1>
- 是否有开源代码: <https://github.com/juyongjiang/BiCAT>

self-supervised learning for sequential recommendation with model augmentation

- 标签:
- 摘要: 这是一篇被ICLR22拒稿的论文。论文提出了三种在模型层面的数据增强策略: 神经元mask、网络层drop和Encoder互补。神经元mask是在Transformer Encoder中的前馈神经网络FFN部分随机mask一些神经元, 网络层drop就是在Encoder之后叠加K层FFN并每次随机drop其中的M层, Encoder互补就是在下一个点击商品任务上预训练一个Encoder, 它生成的embedding与原Encoder相加作为一个增强样本。



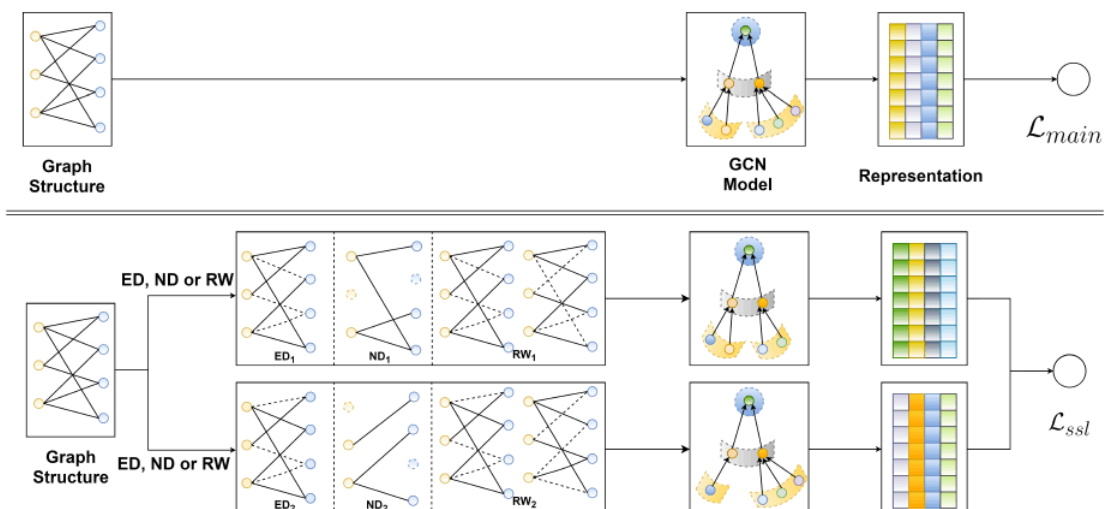
- 链接: <https://openreview.net/forum?id=4YOOO4ZNMK>
- 是否有开源代码: 无

Graph Self-Supervised Learning

图上自监督学习相关论文, 对每篇读过的论文给出基本介绍

Self-supervised Graph Learning for Recommendation[SIGIR'21]

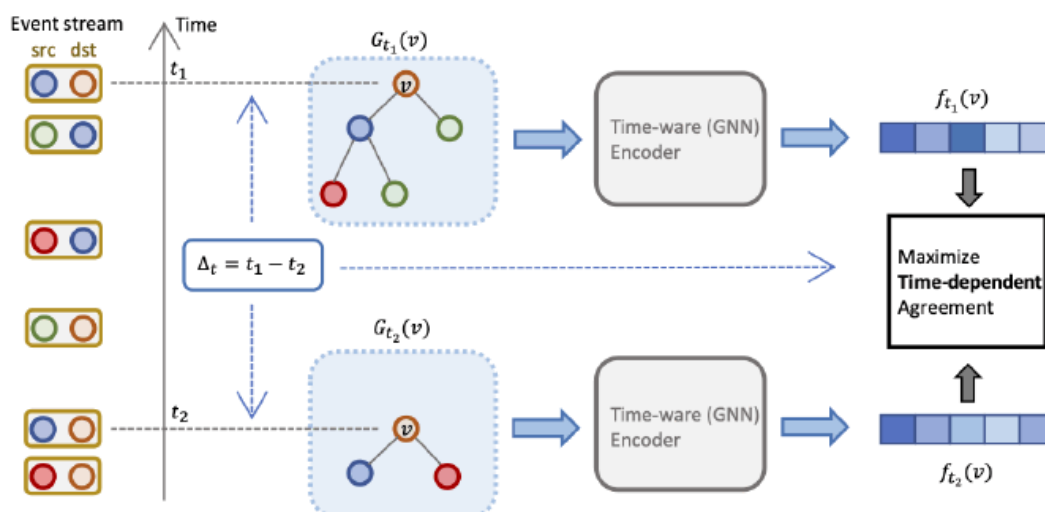
- 标签: 图自监督学习
- 摘要: 在用户-物品交互图上应用对比学习, 提出了图上三种数据增强的方式: 顶点dropout、边dropout以及随机游走, 来生成同一顶点的不同视角, 最小化同一顶点不同视角之间的差异, 同时最大化不同顶点的视角之间的差异。论文花了较大篇幅对加入对比学习导致的性能提升给出了理论分析。



- 链接: <https://dl.acm.org/doi/abs/10.1145/3404835.3462862>
- 是否有开源代码: <https://github.com/wujcan/SGL>

Self-supervised Representation Learning on Dynamic Graphs[CIKM'21]

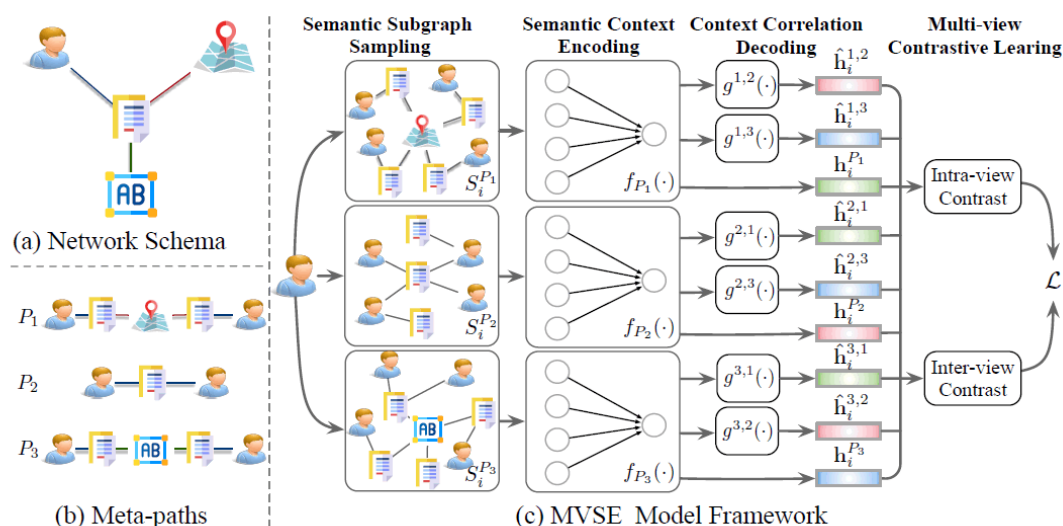
- 标签：图自监督学习、动态图学习
- 摘要：论文提出了一个去偏的动态图对比学习框架DDGCL。对于动态信息的处理，DDGCL从几个方面展开：首先在将子图转化成embedding时，在GNN中加入了时间 t 的embedding。其次，DDGCL从过去一个给定的时间间隔内随机取一个时刻对应的子图作为正样本，认为这段时间内顶点的标签没有发生变化。最后，在进行对比学习时，用了一个时间相关的相似度函数，对时间间隔进行了惩罚，避免两个子图结构相似的顶点在较长的时间间隔下也被判断为相似。



- 链接：<https://dl.acm.org/doi/abs/10.1145/3459637.3482389>
- 是否有开源代码：无

Multi-View Self-Supervised Heterogeneous Graph Embedding[ECML-PKDD'21]

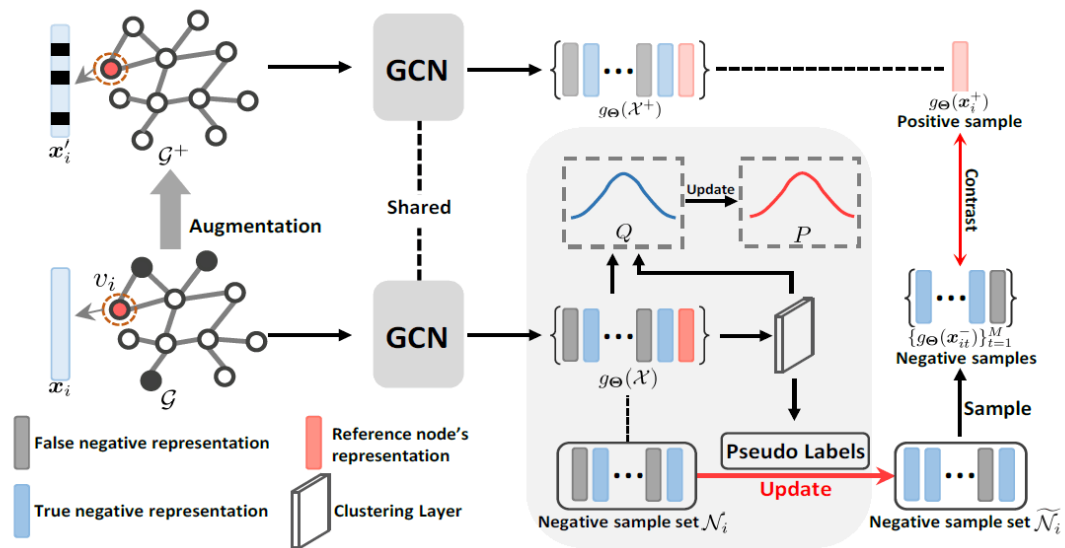
- 标签：元路径、异构图
- 摘要：给定异构图中一个顶点和关联的若干条元路径，论文首先对每条元路径通过固定长度的随机游走来获得元路径相关的子图，并通过GIN来得到每个子图的embedding。因为作者认为不同的元路径代表不同视角，所以顶点所形成的多个子图的embedding就构成了多视角的embedding，作为下一步MoCo对比学习框架输入的key和query。intra-view对比是同一元路径下，不同顶点之间进行对比，inter-view是同一顶点下，不同元路径之间进行对比



- 链接: https://2021.ecmlpkdd.org/wp-content/uploads/2021/07/sub_408.pdf
- 是否有开源代码: <https://github.com/Andy-Border/MVSE>

Graph Debiased Contrastive Learning with Joint Representation Clustering[IJCAI'21]

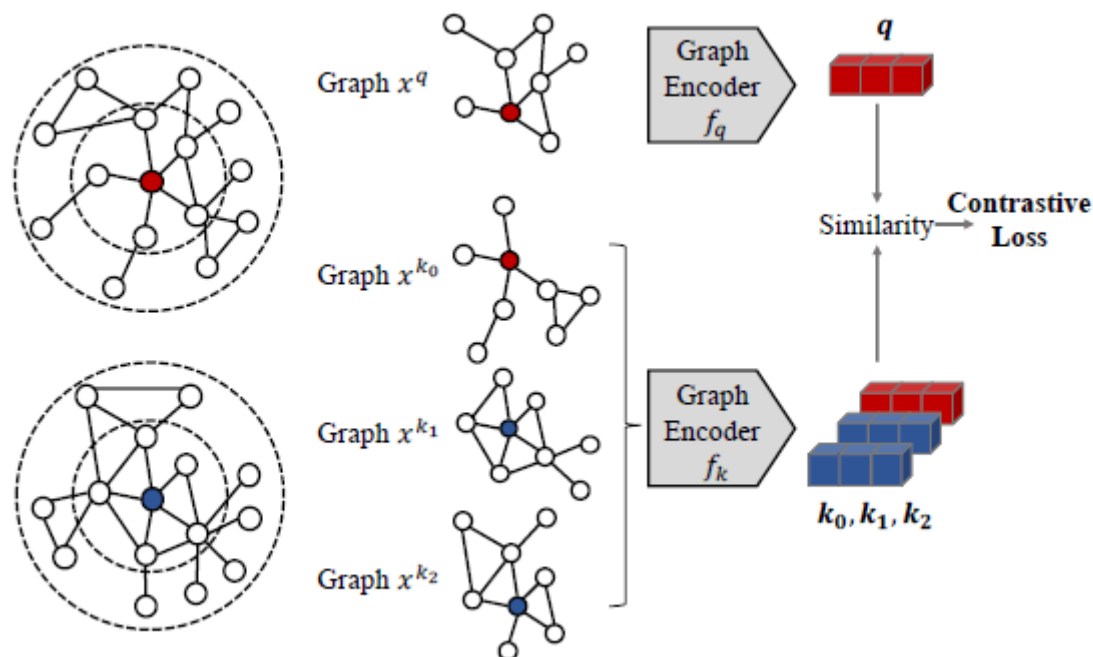
- 标签: 聚类
- 摘要: 现有的对比学习方法大多直接在图上随机采样作为负样本, 这种做法是有偏的因为可能采到假阴性样本从而导致性能下降。论文提出的去偏方法是通过聚类来将不同簇的顶点作为负样本, 减少图上随机采样的偏差。具体地, 正样本是通过将输入顶点的特征随机mask, 而负样本是通过将输入顶点经过GCN得到的embedding聚类后, 选取不同簇的顶点。因此, 聚类结果的好坏会很大影响模型的性能, 论文使用KL散度与t分布来对每次迭代的聚类结果进行评估及更新。



- 链接: <https://www.ijcai.org/proceedings/2021/473>
- 是否有开源代码: <https://github.com/hzhao98/GDCL>

GCC--Graph Contrastive Coding for Graph Neural Network Pre-Training[KDD'20]

- 标签: 随机游走
- 摘要: GCC作为图对比学习的一个方法, 提出了自己的构造正负样本的方式。GCC以子图为单位来构造样本, 这里的子图定义为r中心网络, 即与中心顶点最短距离不大于r的顶点所形成的集合。在构造正负样本时, 给定一个中心顶点, 通过带重启的随机游走->子图抽取->匿名化三个步骤得到它的正样本, 将不同顶点得到的视为负样本。得到的子图样本经过GIN得到各自的embedding, 通过InfoNCE进行对比学习。



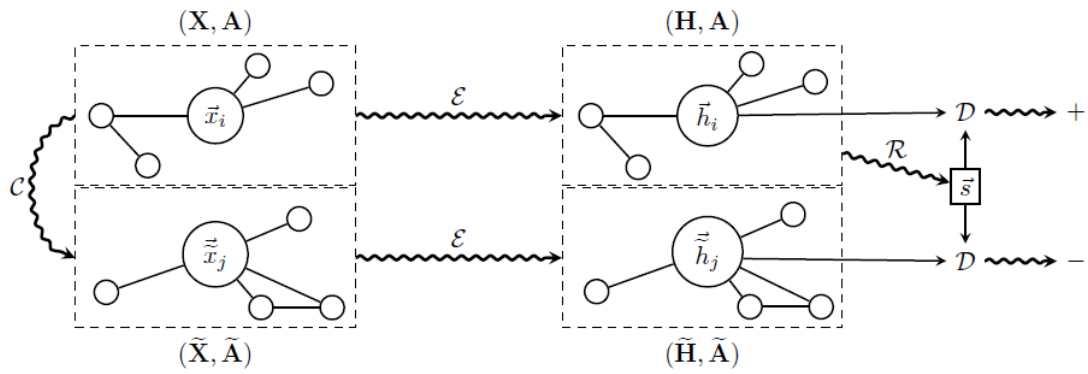
- 链接: <https://www.kdd.org/kdd2020/accepted-papers/view/gcc-graph-contrastive-coding-for-graph-neural-network-pre-training>
- 是否有开源代码: <https://github.com/THUDM/GCC>
- 数据量级: LiveJournal[n=484,3953]

GCCAD: Graph Contrastive Coding for Anomaly Detection

- 标签:
- 摘要:
- 链接: <https://arxiv.org/pdf/2108.07516.pdf>
- 是否有开源代码: <https://github.com/allanchen95/GCCAD>

Deep Graph Infomax[ICLR'19]

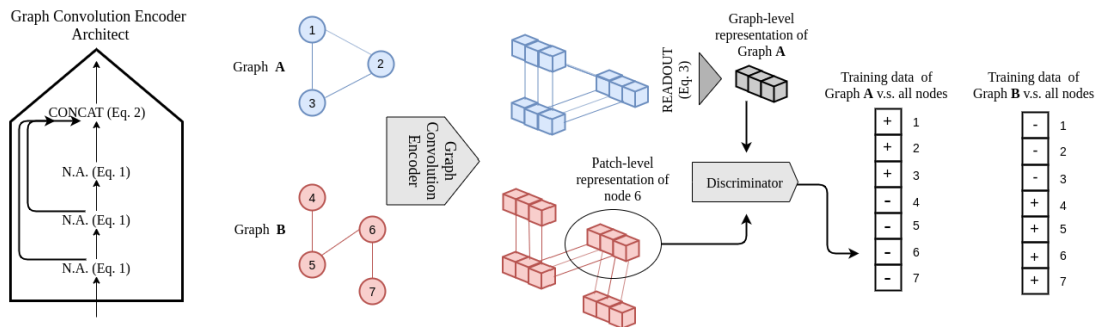
- 标签: 互信息
- 摘要: DGI将Infomax准则用在了图上, 所谓的Infomax准则是指最大化系统输入与输出之间的互信息。DGI通过最大化顶点与图特征表示之间的互信息来进行对比学习。输入一张图, DGI通过一个编码器例如GCN得到图中顶点的特征表示 H 后, 利用一个readout函数例如average pooling聚合顶点的特征表示得到图的特征表示 S 。顶点和图的特征表示彼此构成正样本对, 而对原图进行扰乱后经过相同编码器得到的顶点的特征表示 H' 表示则与原图的特征表示 S' 构成负样本对。按照Infomax准则, DGI中的系统指的是readout函数, 输入为顶点特征 H 输出为图的特征表示 S 。用于区分正负样本的判别器选取双线性评分函数 $D(h_i, s) = \sigma(h_i^T W s)$, 输出的是 (h_i, s) 为正样本对的概率, 用于更新特征表示的损失函数是Jensen-Shannon散度。关于DGI更详细的解析可见:
 - [互信息及其在图表示学习中的应用](#)
 - [深度学习中的互信息: 无监督提取特征](#)



- 链接: <https://openreview.net/forum?id=rklz9iAckQ>
- 有无开源代码: <https://github.com/PetarV-/DGI>
- 数据量级: Reddit[n=23,1443]

InfoGraph: Unsupervised and Semi-supervised Graph-Level Representation Learning via Mutual Information Maximization[ICLR'20]

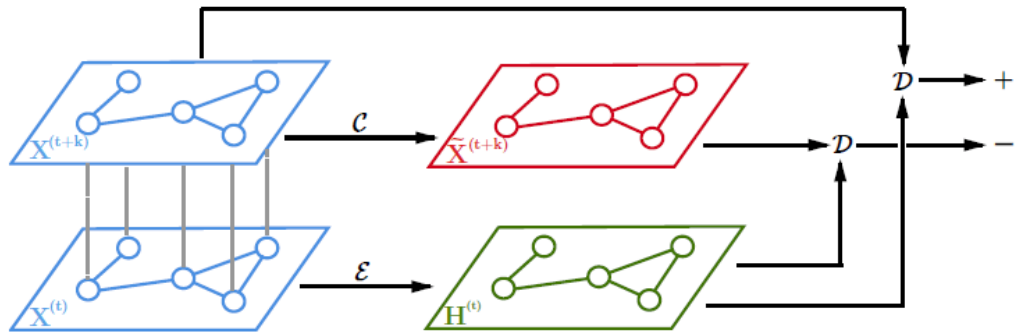
- 标签: 互信息
- 摘要: InfoGraph的输入是一系列图, 无监督任务下为每个图学习一个特征表示, 半监督任务下利用给定的有限带标签图为未知图预测。InfoGraph的无监督学习部分与DGI类似, 编码器上它使用GIN而不是DGI中的GCN, 将 K 层后GIN得到的各顶点的特征表示通过readout函数聚合后得到图的特征表示 S , 而对GIN每一层得到的顶点特征表示进行拼接, 得到patch特征表示作为正样本 H 。 S 表示全局信息而 H 表示局部信息。在负样本的选择上, InfoGraph将当前图的patch与不同图的特征表示进行组合作为负样本。在半监督任务下, 为了更好地进行知识迁移, 同时最大化监督部分的编码器与无监督部分的编码器之间的互信息。



- 链接: <https://arxiv.org/abs/1908.01000>
- 有无开源代码: <https://github.com/fanyun-sun/InfoGraph>
- 数据量级: RDT-M5K[n=4999*508.52]

Spatio-Temporal Deep Graph Infomax[ICLR'19 workshop]

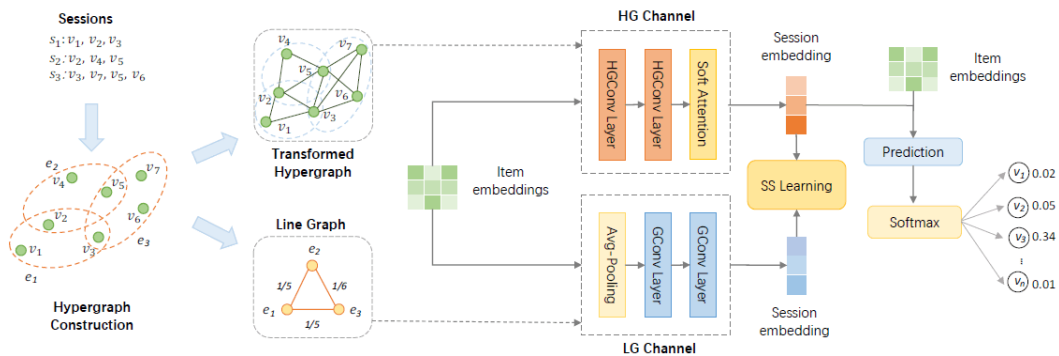
- 标签: 互信息、动态图学习
- 摘要: STDGI在DGI的基础上加入动态的时间信息, 体现在正负样本的构造上: 对于 t 时刻的图上所有顶点, 通过一个编码器如GCN得到它们的embedding $H^{(t)}$, 与 $t+k$ 时刻的顶点特征 $X^{(t+k)}$ 构成正样本; 与 $t+k$ 时刻随机mask的顶点特征构成负样本



- 链接: <https://rlgm.github.io/papers/48.pdf>
- 有无开源代码: 无
- 数据量级: METR-LA[n=207]

Self-Supervised Hypergraph Convolutional Networks for Session-based Recommendation[AAAI'21]

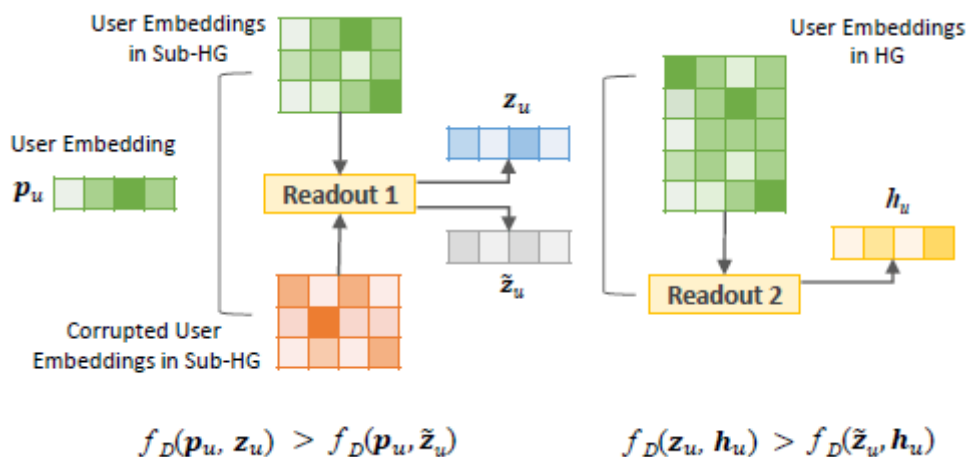
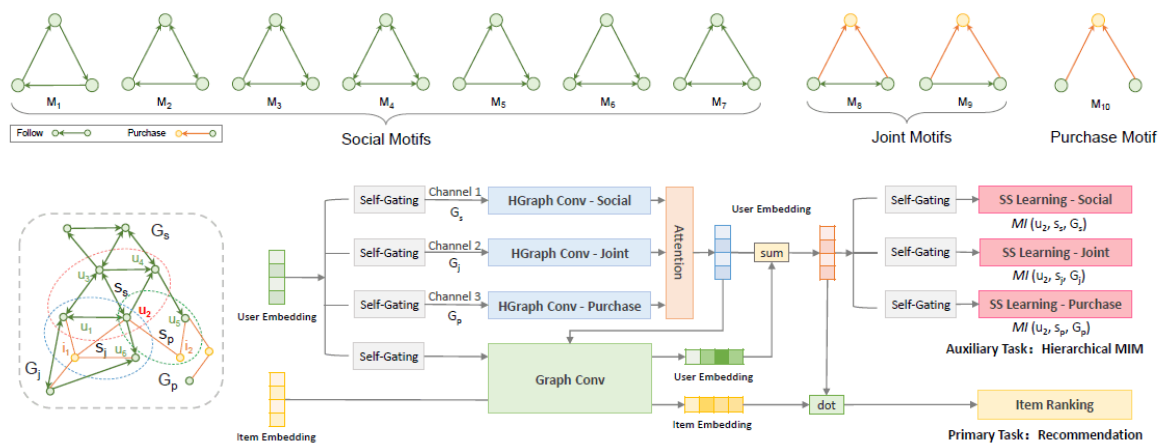
- 标签: 超图学习
- 摘要: 论文在序列推荐任务上提出了DHCN方法, 结合了超图学习和图对比学习的理论。所谓的超图是指将子图看作顶点, 在子图之间添加边, 构成一张超图。DHCN将一条用户点击序列看作一个子图, 序列内任意两个物品之间构成一条边。这样构成了下图中的"Transformed Hypergraph", 以物品为顶点; 而序列之间共享的物品作为连接两个子图的边。构成了下图中的"Line Graph", 以子图为顶点。两个图就构成了同一条序列的两个视角, 通过图卷积得到各自的embedding后, 同一序列的两个视角构成正样本, 不同序列视为负样本。判别器选择将两个输入的embedding点乘的形式, 损失函数为InfoNCE。



- 链接: <https://ojs.aaai.org/index.php/AAAI/article/view/16578/16385>
- 有无开源代码: <https://github.com/xiaxin1998/DHCN>
- 数据量级: Nowplaying[n=91,5128]

Self-Supervised Multi-Channel Hypergraph Convolutional Network for Social Recommendation[WWW'21]

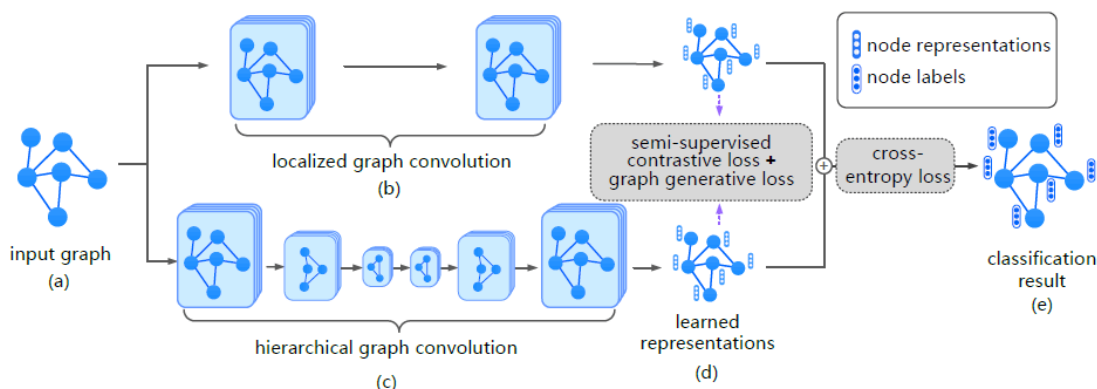
- 标签: 超图学习、Motif
- 摘要: 论文在社会推荐任务上提出了MHCN方法, 根据社会推荐场景中的好友-好友、好友-商品、用户-商品三种典型的关系模式, 通过三类十种三角Motif来构造三个视角的超图进行对比学习。不同于DGI直接进行顶点-图的互信息最大化, MHCN加入了以顶点为中心的子图来进行顶点-子图-图的层次化互信息最大化。构造正负样本时, 将子图中顶点的embedding顺序随机打乱来得到扰乱子图, 作为顶点-子图和子图-图两个阶段对比学习的负样本, 判别器选取的是将输入的两个embedding进行点乘, 损失函数不同于常见的交叉熵选取了排序损失。



- 链接: <https://dl.acm.org/doi/fullHtml/10.1145/3442381.3449844>
- 有无开源代码: <https://github.com/Coder-Yu/RecQ>
- 数据量级: Yelp[n=1,9539]

Contrastive and Generative Graph Convolutional Networks for Graph-based Semi-Supervised Learning[AAAI'21]

- 标签:
- 摘要: 论文解决的是如何利用有限带标签的样本来为大量无标签样本预测。分为无监督和有监督任务，从局部和全局两个视角进行对比学习，GCN因任务为是对邻域信息聚合，所以作为局部视角下的编码器，而Hierarchical GCN用来产生全局视角。在无监督任务下，因为没有数据的标签信息，论文将同一顶点的两个视角看作正样本，不同顶点视为负样本。在有监督任务时，同一类的为正样本而不同类的为负样本。

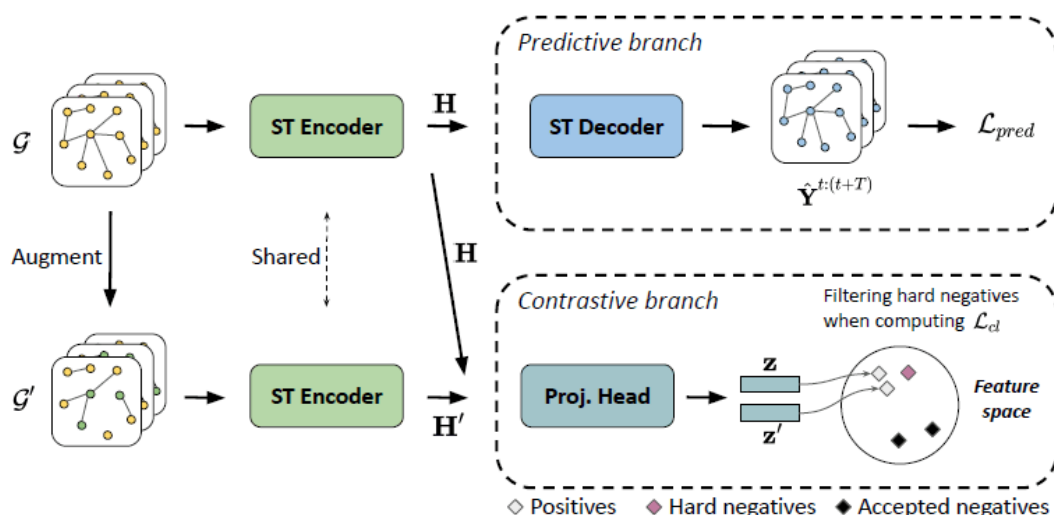


- 链接: <https://ojs.aaai.org/index.php/AAAI/article/view/17206/17013>
- 有无开源代码: <https://github.com/LEAP-WS/CG3>

- 数据量级：PubMed[n=1,9717]

Spatio-Temporal Graph Contrastive Learning

- 标签：动态图学习
- 摘要：STGCL的输入是传感器的采集数据，在构建正样本时，除了传统的特征mask和边dropout，还对相邻时刻的特征进行线性插值；在构建负样本时，为了避免采样到临近时刻的“hard negative”样本，论文只考虑时间间隔大于给定阈值的负样本。



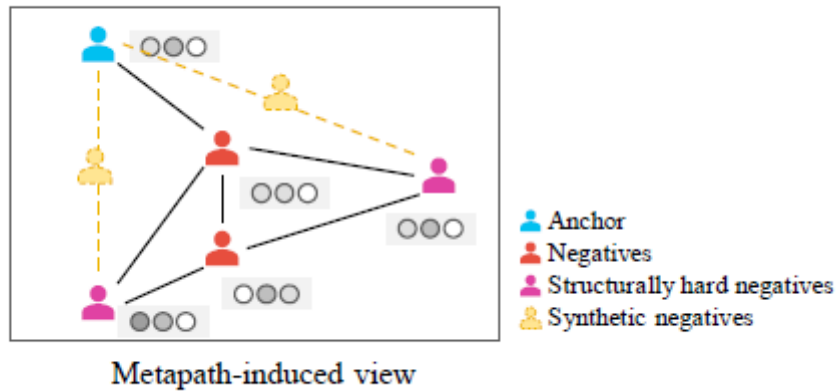
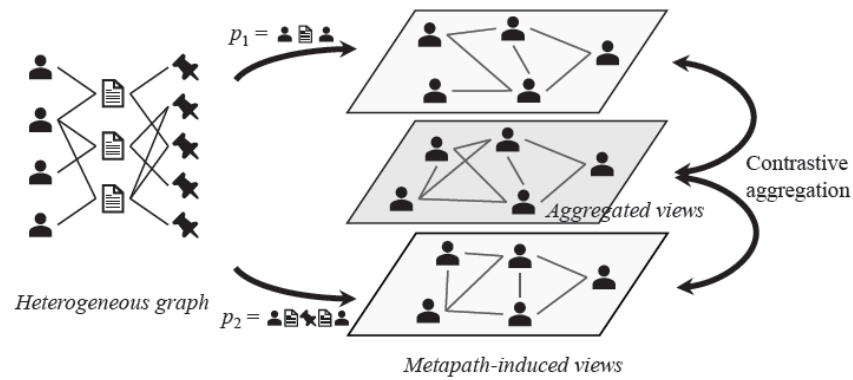
- 链接：<https://arxiv.org/abs/2102.12380>
- 有无开源代码：<https://github.com/Mobzhang/PT-DGNN>

TCL: Transformer-based Dynamic Graph Modelling via Contrastive Learning

- 标签：动态图学习
- 摘要：TCL的目标是利用历史信息预测 t 时刻用户是否会与物品产生交互，使用动态依赖-交互图作为图的呈现，各自得到用户和物品的 k 最短距离子图后，通过BFS转化为序列后输入定制的Transformer中，得到各自的特征表示。在构建正负样本时，用户已交互过的物品作为正样本，其余未交互过的物品作为负样本
- 链接：<https://arxiv.org/abs/2105.07944>
- 有无开源代码：无

Structure-Enhanced Heterogeneous Graph Contrastive Learning

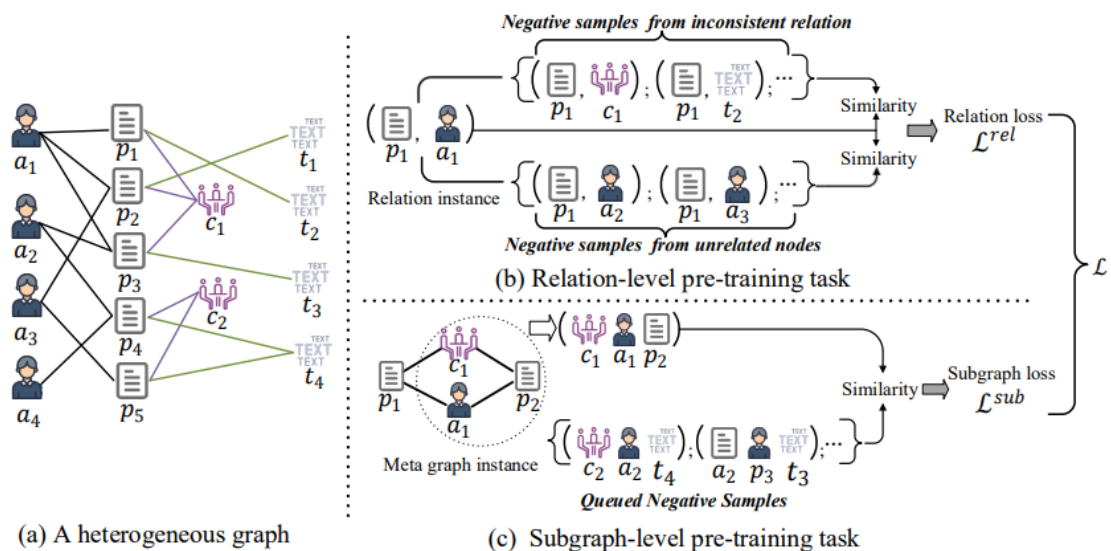
- 标签：异构图
- 摘要：一个异构图可以根据不同的元路径切分为多个子图，顶点在每个子图下的特征表示作为局部视角，在全下的特征特征表示作为全局视角，一个顶点的局部与全局视角之间为正样本，与其他顶点的局部、全局视角分别构成负样本。在负样本的选取上，进一步提出了根据Top-K的PPR值来选取hard negative样本，同时通过线性加权引入合成负样本。



- 链接: <https://www.cs.emory.edu/~jyang71/files/stencil.pdf>
- 是否有开源代码: <https://github.com/CRIPACDIG/STENCIL>

Contrastive Pre-Training of GNNs on Heterogeneous Graphs[CIKM'21]

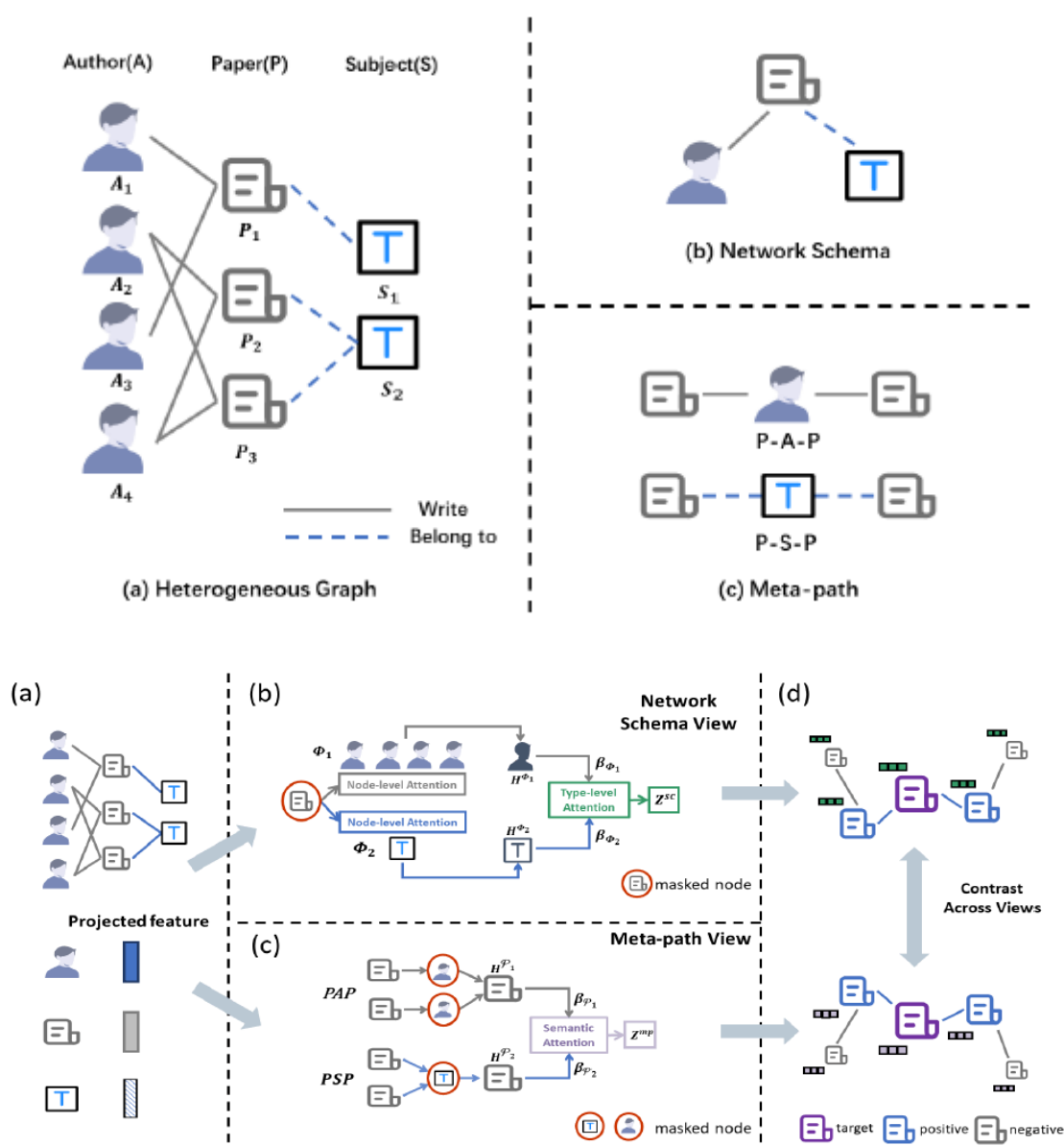
- 标签: 异构图
- 摘要: 论文在边和子图两个层面提出了自监督的预训练任务: 边预训练任务中, 对于一个顶点, 不同类型的边相连的顶点为一种负样本, k跳邻居为另一种负样本; 子图预训练任务中, 对于一个顶点, 通过元路径形成的元图中的其他顶点为正样本, 上个batch中的正样本为当前batch的负样本。



- 链接: <https://yuanfulu.github.io/publication/CIKM-CPT.pdf>
- 是否有开源代码: 无

Self-supervised Heterogeneous Graph Neural Network with Co-contrastive Learning[KDD'21]

- 标签：元路径、attention、异构图
- 摘要：HeCo通过一个异构图的元路径和网络结构两个角度来学习顶点的embedding。网络结构视角通过node-level和type-level两层attention，前者在特定的顶点类型下聚合邻域信息，后者再对不同的顶点类型的结果进行聚合；元路径视角通过元路径定义一个顶点的邻域，在每种元路径下通过GCN聚合邻域信息，再对不同元路径的结果通过attention进行聚合。论文的对比方式是，它认为有多条元路径相连的两个顶点构成一对正样本，而元路径数量少于设定阈值的两个顶点视为一对负样本。同时，原样本的表示取的是网络结构视角，而正负样本的表示取的是元路径视角。



- 链接: <https://arxiv.org/abs/2105.09111>
- 是否有开源代码: <https://github.com/liun-online/HeCo>