

Beijing Forest Studio
北京理工大学信息系统及安全对抗实验中心



时序网络嵌入方法介绍

李新帅

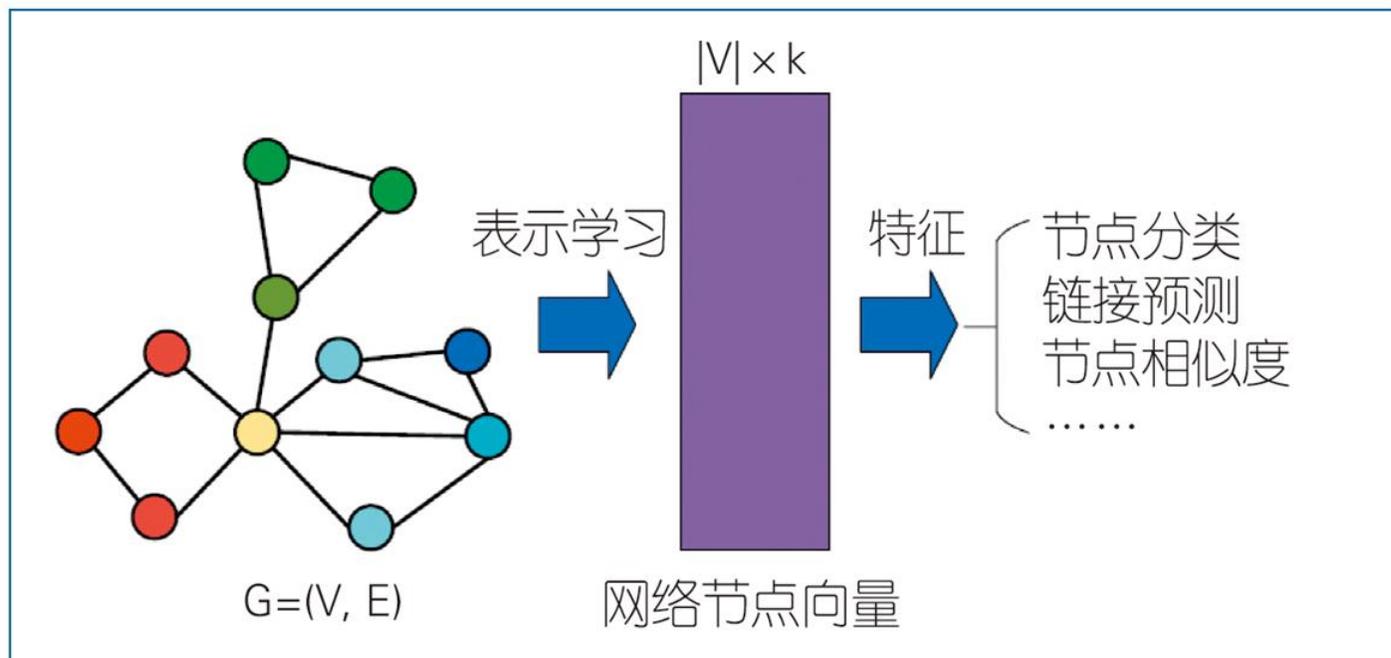
导师：潘丽敏

2020年11月15日

- 背景简介
- 基本概念
- 算法原理
- 优劣分析
- 应用总结
- 参考文献

- 预期收获
 - 1.了解网络嵌入基本思想
 - 2.了解动态网络嵌入的学习方法
 - 3.了解动态网络嵌入的应用

- 网络嵌入 (Network Embedding)
 - 网络表示学习 (Network Representation Learning)
 - 图嵌入 (Graph Embedding)

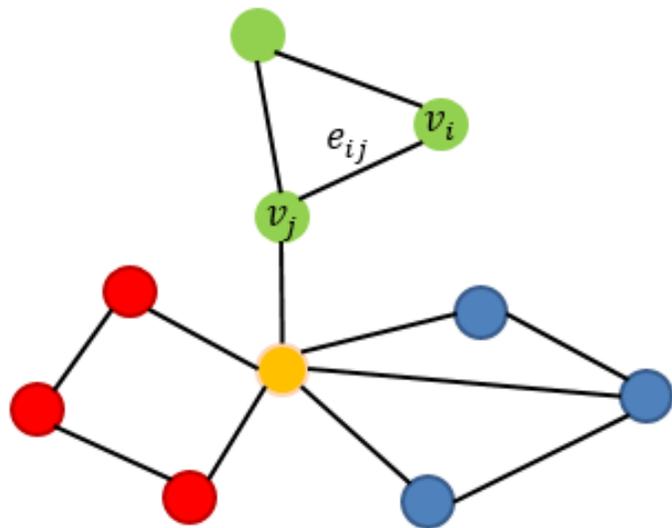




基本概念

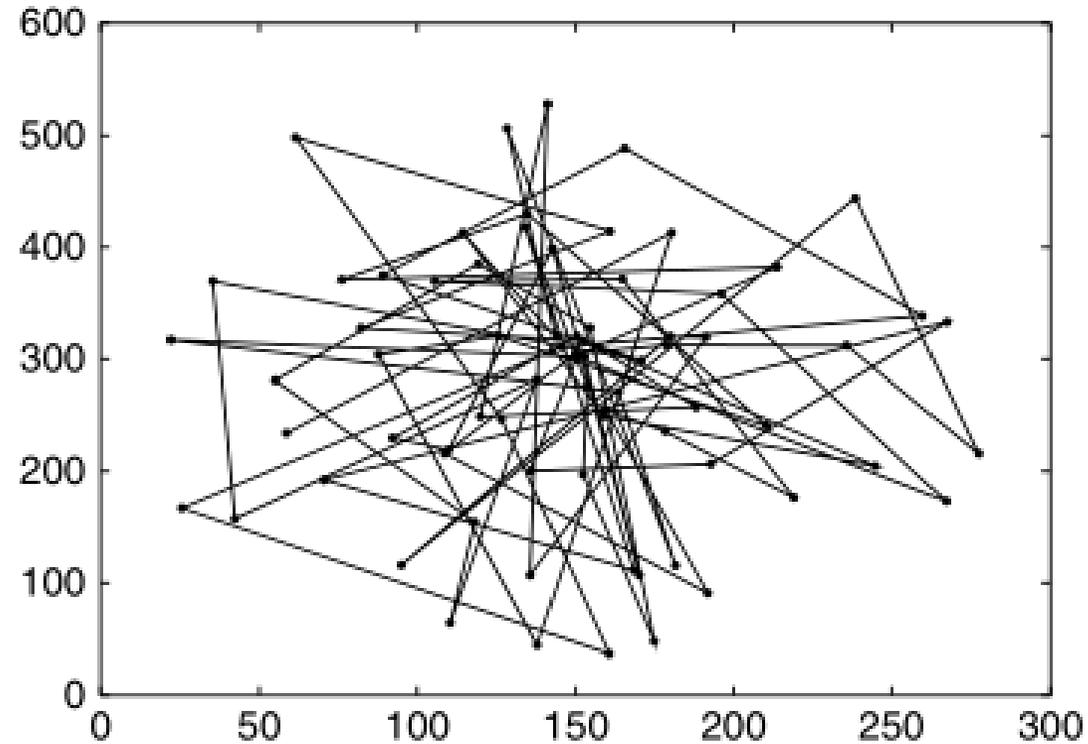
- 图

– 将网络记为图 $G(V, E)$ ，其中 $V = \{v_1, v_2, \dots, v_n\}$ 是节点集合， $E = \{e_{ij}\}_{i,j=1}^n$ 是边的集合， e_{ij} 表示节点 v_i 和 v_j 之间的边。



$$G = (V, E)$$

- 随机游走 (random walk)
 - 基于过去的表现，无法预测将来的发展步骤和方向。

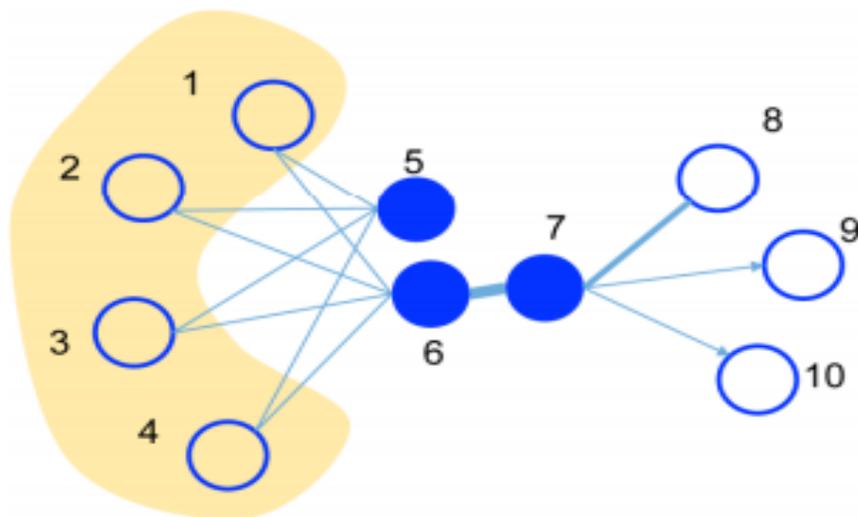


- 一阶相似度

- 用于描述图中成对顶点之间的局部相似度，形式化描述为若 v_i, v_j 之间存在直连边，则边权 $v_{i,j}$ 即为两个顶点的相似度，若不存在直连边，则1阶相似度为0。

- 二阶相似度

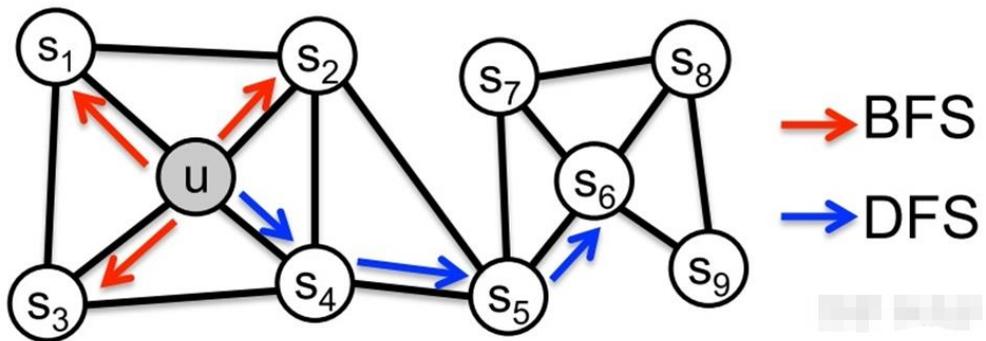
- 二阶相似度描述了一对节点的邻域结构的接近程度。即两个节点相同的用户越多，这两个节点越相似。



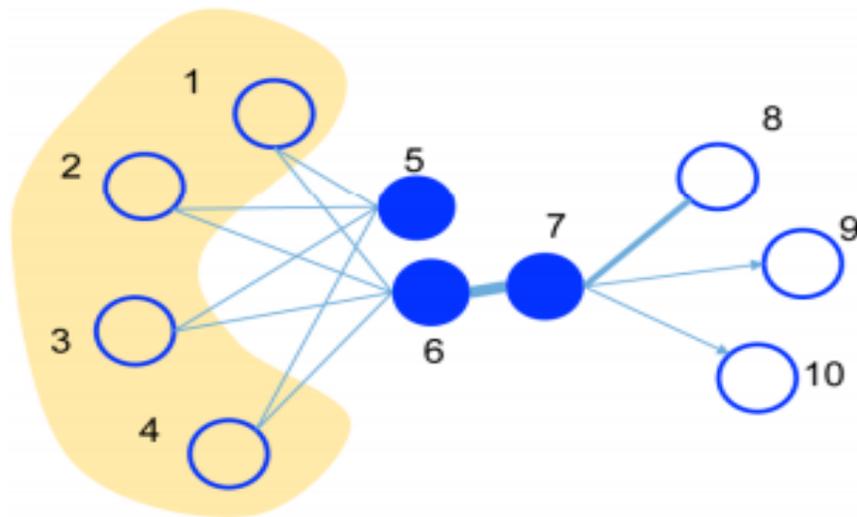
- 时序点过程（Temporal point process）
 - 是由连续时间域上的一系列（带标值）事件构成的随机过程。
 - 核心是其条件强度函数，有时可以简称为强度函数，用 $\lambda^*(t)$ 表示。
 - 考虑所有历史事件的影响来实现对历史信息的充分利用。
- 霍克斯过程（Hawkes processes）
 - 自激励过程，其内部机制表示发生的历史事件对于未来事件的发生有激励作用，并且历史事件的影响以累加的形式进行叠加。

$$\lambda^*(t) = \lambda_0 + \sum_{t_i < t} g_{exc}(t - t_i)$$

- 静态网络嵌入表示
 - Deepwalk (2014)
 - Line (2015)
 - Node2vec (2016)



Node2vec



Line

- 动态网络嵌入表示——离散型表示
 - 使用一系列动态网络的**快照**来表示网络
 - 基于矩阵分解 (Zhu et.al, 2016)
 - DynGEM (Goyal et al, 2018)
- 动态网络嵌入表示——连续型表示
 - 能够记录动态网络完整的时序信息
 - DySAT (Sankar et al, 2018)
 - **HTNE** (Zuo et al, 2018)



算法原理

T	对网络完整地时序形成过程进行学习建模，将网络中的节点表示成低维、实值、稠密的向量形式
I	网络拓扑结构
P	1.邻居生成序列 2.霍克斯过程，注意力机制 3.参数优化
O	节点向量表示

P	尽可能利用图的拓扑结构信息、时序信息，研究网络的形成过程
C	具备网络拓扑结构、时序信息
D	1.条件强度函数的选择 2.邻居生成序列
L	KDD2018

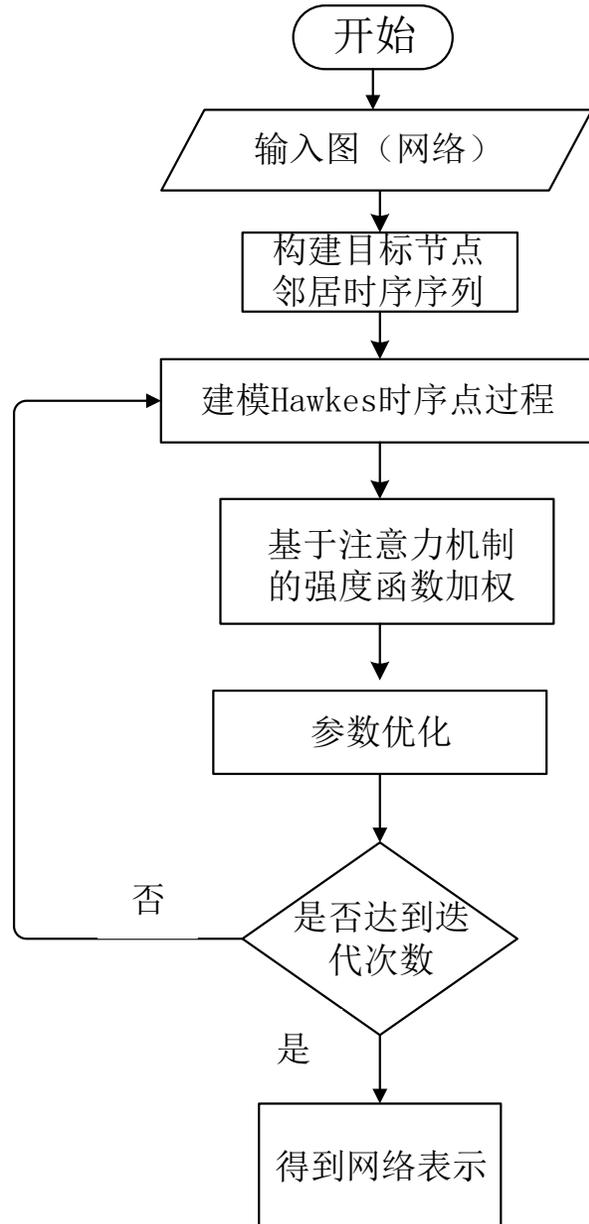
- 提出问题

- 现有的网络嵌入算法普遍关注的是**静态网络结构**，但他们均假设一个节点的邻域是**无序的**，也就是说链路的**形成历史**被忽视。

- 解决办法

- 提出了基于**霍克斯过程**的**时序网络嵌入(HTNE)**算法。具体来说，从**连续事件驱动**的网络结构中归纳了**邻居生成序列**。

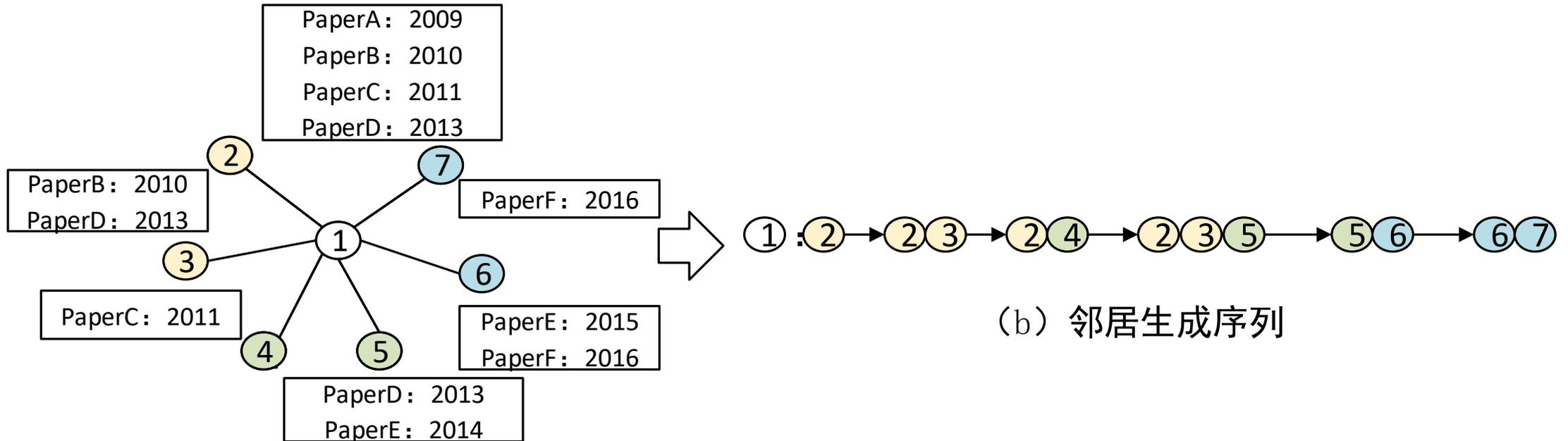
- 算法原理图



- 时序网络 (Temporal Network)
 - 边被节点间依序交互事件所注释的网络
 - 定义为—— $G = (V, E, A)$
 - $a_{x,y} = \{a_1 \rightarrow a_2 \rightarrow \dots\} \in A$
 - 网络中一个节点临近的邻居，可以按节点与其交互事件的发生时间组成一个序列，来表示邻居生成过程

- 邻居生成序列 (Neighborhood Formation Sequence)

- 给出一个时序网络中的源节点 x 属于 V ，该节点的邻居集记为 $N(x)$ ，节点和每个邻居之间的边都用依时间顺序的交互事件所注释。



(a) 共同作者时序网络

(b) 邻居生成序列

- 霍克斯过程 (Hawkes Process)
 - 当前事件的产生不仅仅取决于上一步时间的事件，还受到效果随时间衰减的历史事件影响。

$$\tilde{\lambda}_{y|x}(t) = \mu_{x,y} + \sum_{t_h < t} \alpha_{h,y} K(t - t_h)$$

$$\mu_{x,y} = f(e_x, e_y) = -\|e_x - e_y\|^2$$

$$\alpha_{h,y} = f(e_h, e_y) = -\|e_h - e_y\|^2$$

- 注意力机制 (Attention)
 - 考虑源节点与历史节点之间的影响。

$$w_{h,x} = \frac{\exp(-\|e_x - e_h\|^2)}{\sum_{h'} \exp(-\|e_x - e_{h'}\|^2)}$$

$$\alpha_{h,y} = w_{h,x} f(e_h, e_y)$$

- 条件概率函数

$$\lambda_{y|x}(t) = \exp(\tilde{\lambda}_{y|x}(t))$$

$$p(y|x, \mathcal{H}_x(t)) = \frac{\lambda_{y|x}(t)}{\sum_{y'} \lambda_{y'|x}(t)}$$

$$\log \mathcal{L} = \sum_{x \in \mathcal{V}} \sum_{y \in \mathcal{H}_x} \log p(y|x, \mathcal{H}_x(t))$$

- 目标函数

$$\log \sigma(\tilde{\lambda}_{y|x}(t)) + \sum_{k=1}^K \mathbb{E}_{\mathcal{V}^k \sim P_n(\mathcal{V})} [-\log \sigma(\tilde{\lambda}_{\mathcal{V}^k|x}(t))]$$

- 模型优化

- 使用**随机梯度下降**（SGD）方法，对模型参数进行优化。
- 每轮迭代中，随机优化某一条训练数据上的损失函数，这样每一轮参数的更新速度大大加快。
- 为了降低模型**复杂度**，规定历史节点的最大长度，逐步优化模型参数。

- 数据集
 - DBLP, Yelp, Tmall
- 对比方法
 - DeepWalk, LINE, node2vec, ComE
- 对比实验：
 - 节点分类, 链路预测, 可视化任务

• 节点分类

Metric	Method	10%	20%	30%	40%	50%	60%	70%	80%	90%
Macro-F1	DeepWalk	0.6345	0.6553	0.6635	0.6681	0.6698	0.6721	0.6734	0.6725	0.6745
	node2vec	0.6332	0.6511	0.6589	0.6631	0.6655	0.6667	0.6670	0.6639	0.6660
	LINE	0.6163	0.6350	0.6415	0.6455	0.6474	0.6489	0.6498	0.6466	0.6490
	ComE	0.6508	0.6632	0.6680	0.6718	0.6753	0.6764	0.6794	0.6769	0.6791
	HTNE	0.6235	0.6409	0.6490	0.6526	0.6564	0.6592	0.6596	0.6570	0.6608
	HTNE-a	0.6528	0.6656	0.6729	0.6768	0.6799	0.6824	0.6854	0.6836	0.6844
Micro-F1	DeepWalk	0.6435	0.6604	0.6662	0.6690	0.6700	0.6711	0.6711	0.6709	0.6719
	node2vec	0.6492	0.6626	0.6688	0.6717	0.6736	0.6742	0.6742	0.6730	0.6735
	LINE	0.6229	0.6371	0.6433	0.6455	0.6476	0.6484	0.6487	0.6470	0.6463
	ComE	0.6608	0.6723	0.6758	0.6782	0.6806	0.6810	0.6816	0.6801	0.6816
	HTNE	0.6620	0.6693	0.6737	0.6752	0.6778	0.6797	0.6793	0.6777	0.6784
	HTNE-a	0.6706	0.6789	0.6834	0.6853	0.6869	0.6881	0.6883	0.6879	0.6866

DBLP数据集节点分类结果

• 节点分类

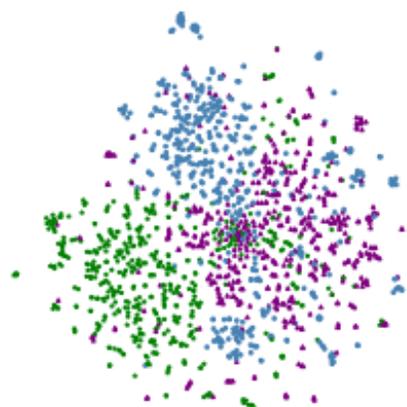
Metric	Method	10%	20%	30%	40%	50%	60%	70%	80%	90%
Macro-F1	DeepWalk	0.4862	0.4892	0.4913	0.4922	0.4923	0.4927	0.4939	0.4941	0.4940
	node2vec	0.5298	0.5348	0.5363	0.5377	0.5368	0.5376	0.5386	0.5391	0.5391
	LINE	0.4311	0.4350	0.4364	0.4370	0.4370	0.4369	0.4382	0.4387	0.4384
	ComE	0.5373	0.5416	0.5435	0.5442	0.5442	0.5451	0.5465	0.5455	0.5428
	HTNE	0.5292	0.5413	0.5476	0.5511	0.5524	0.5539	0.5559	0.5563	0.5563
	HTNE-a	0.5373	0.5433	0.5468	0.5479	0.5485	0.5491	0.5496	0.5493	0.5507
Micro-F1	DeepWalk	0.5652	0.5704	0.5721	0.5732	0.5736	0.5742	0.5749	0.5759	0.5758
	node2vec	0.5971	0.6025	0.6037	0.6049	0.6046	0.6052	0.6059	0.6068	0.6067
	LINE	0.5285	0.5339	0.5358	0.5365	0.5369	0.5370	0.5377	0.5388	0.5388
	ComE	0.6059	0.6100	0.6110	0.6117	0.6119	0.6126	0.6136	0.6131	0.6113
	HTNE	0.6219	0.6286	0.6314	0.6328	0.6332	0.6339	0.6345	0.6352	0.6343
	HTNE-a	0.6194	0.6231	0.6248	0.6251	0.6253	0.6259	0.6259	0.6262	0.6266

Tmall数据集节点分类结果

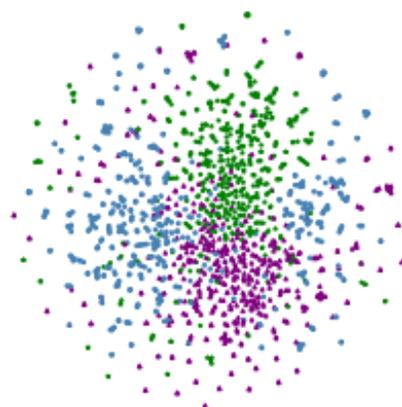
- 链路预测

	DBLP	Yelp	Tmall
DeepWalk	0.8126	0.7678	0.7745
LINE	0.6350	0.8529	0.8265
node2vec	0.8049	0.7712	0.5901
ComE	0.7921	0.8120	0.6917
HTNE	<u>0.8521</u>	0.8944	0.7834
HTNE-a	0.8608	<u>0.8861</u>	<u>0.7928</u>

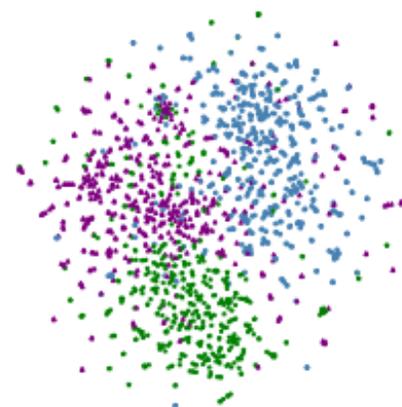
- 可视化任务



(a) LINE



(b) DeepWalk



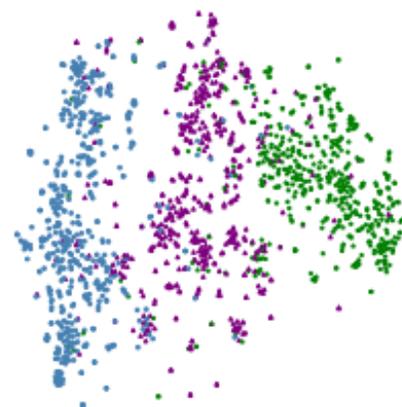
(c) node2vec



(d) ComE



(e) HTNE



(f) HTNE-a



优劣分析

- **优势**
 - 考虑网络邻居的生成过程，捕获了更多有意义的信息，丰富的辅助信息来帮助节点嵌入表示
 - 邻居节点可以在序列中反复出现
- **劣势**
 - 对数据集中时序信息的要求过高



应用总结

- 算法的应用领域
 - 节点分类
 - 链接预测
 - 可视化任务
 - 社区发现
- 未来发展
 - 大规模网络的网络嵌入
 - 异构网络的网络嵌入
 - ...

- [1] 涂存超, 杨成, 刘知远,等. 网络表示学习综述[J]. 中国科学:信息科学, 2017(8).
- [2] Zuo Y , Liu G , Lin H , et al. Embedding Temporal Network via Neighborhood Formation[C]// 2018:2857–2866. KDD2018.
- [3] Peng C , Xiao W , Jian P , et al. A Survey on Network Embedding[J]. IEEE Transactions on Knowledge and Data Engineering, 2017, PP(99):1–1.
- [4] Lu Y , Wang X , Shi C , et al. Temporal Network Embedding with Micro- and Macro-dynamics[J]. CIKM2019.

知人者智，自知者明。胜人者有力，自胜者强。知足者富。强行者有志。不失其所者久。死而不亡者，寿。

谢谢！

