异步社区
人民邮电出版社

图表示学习1.0

讲师：Houye

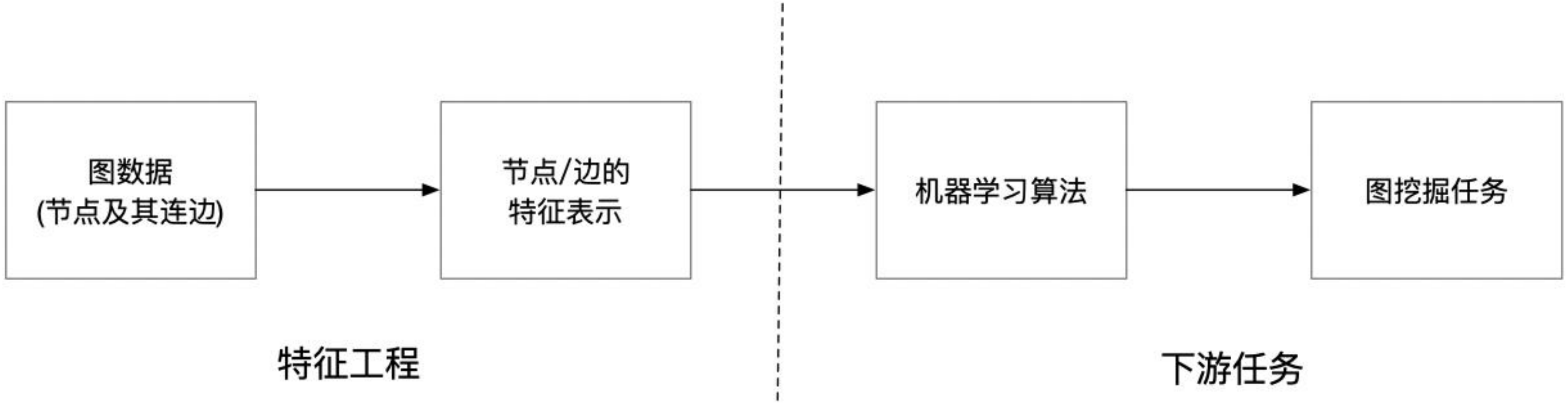
目录

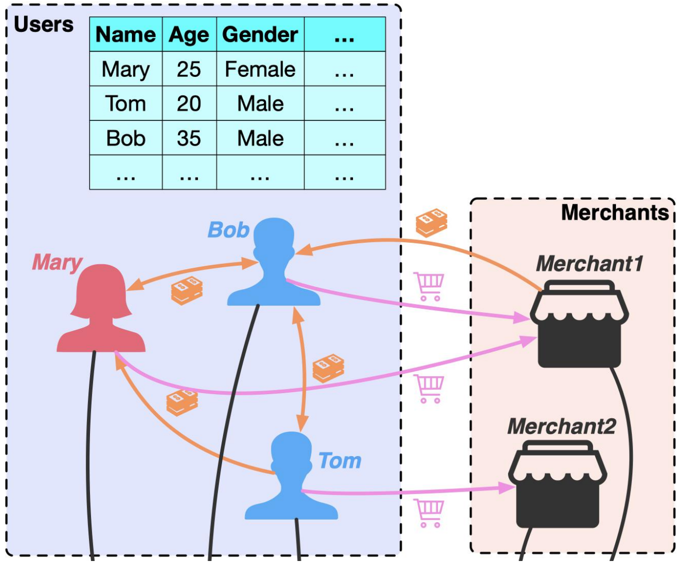
CONTENTS

01 图表示学习

图表示学习

图机器学习流程





- 推荐(链路预测)
- 异常检测(边分类)
- 社区检测(聚类)

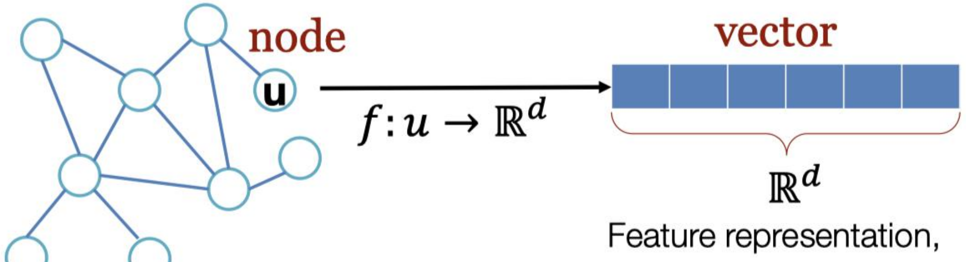
图表示学习

图上的手动特征工程

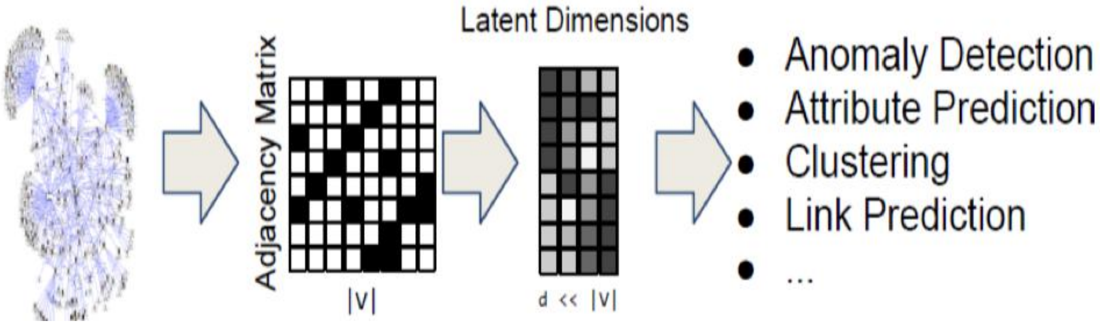
- 抽取节点的特征.例如,社交网络中用户的身高,体重,性别.
- 边也是一样.

图表示学习(图上的自动特征工程)

- 从图数据中自动学习节点/边的特征表示(一个d维向量).



Name	Age	Gender	...
Mary	25	Female	...
Tom	20	Male	...
Bob	35	Male	...
...	...	...	...



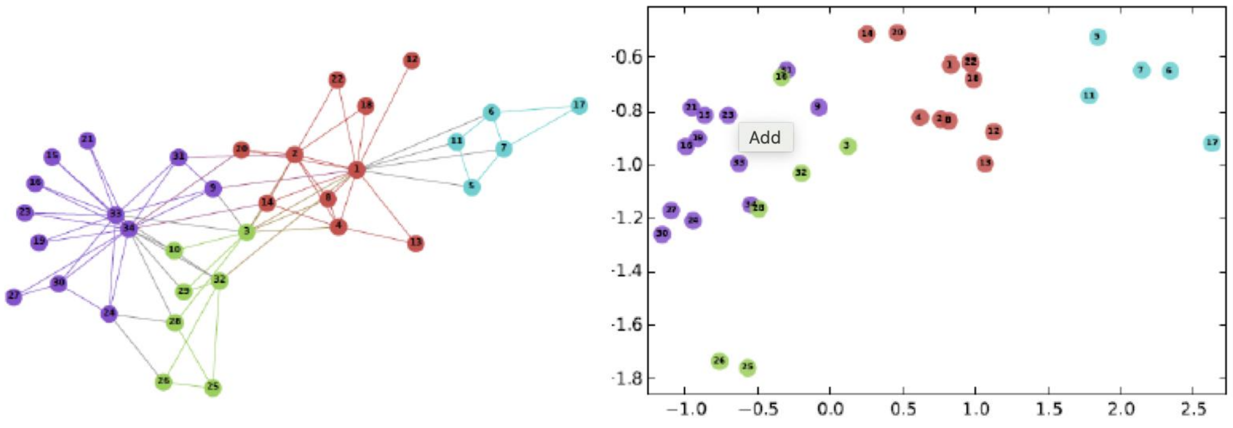
Graph Embedding, Graph Representation Learning
Network Embedding, Network Representation Learning

图上现有的特征(节点属性) -> 基于图结构自动学习特征表示(节点Embedding)

图表示学习

为什么图表示学习?

- 手动特征工程主要考虑了节点的属性信息,忽略了节点的结构信息.
  - 如果节点没有属性, 该怎么办?
- 可以保持图本身的结构和性质,这是图数据的本质特点.
- 可以自动生成节点的表示,避免了手动特征工程带来的偏差.
  - 哪些特征的是有效特征?

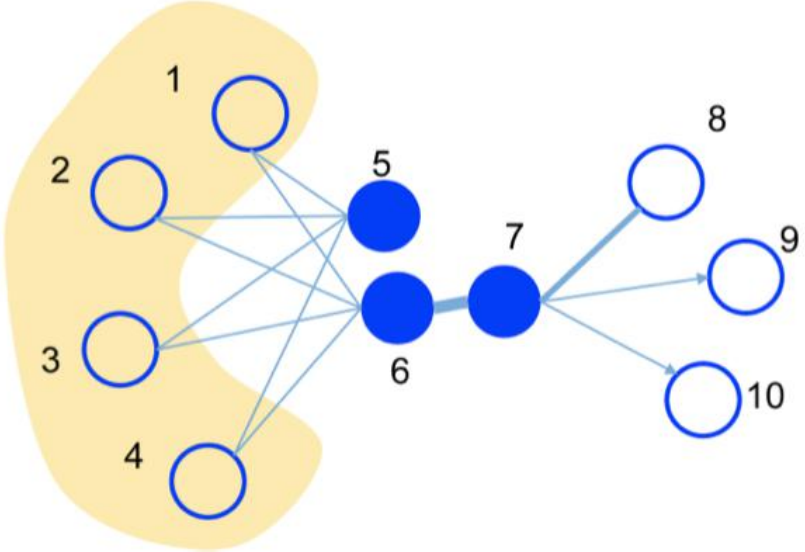


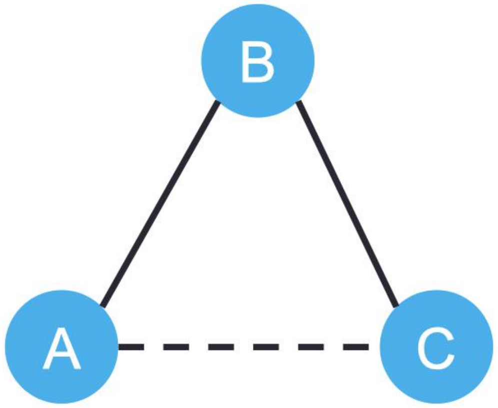
Input　　Output

图表示学习

图表示学习的要求

- 保持图数据本身的**结构和性质**.
- 图的结构信息(节点间的连接情况)可以反映节点的相似性.
  - 连边越多,边权重越大,节点的相似性越高.
  - 一阶结构, 二阶结构等反映不同的结构信息(节点相似性)
- 图数据本身的性质 -> 节点的特性.
  - 传递性,非对称性.





Transitivity

Dist(A,B)+Dist(B,C)>Dist(A,C)

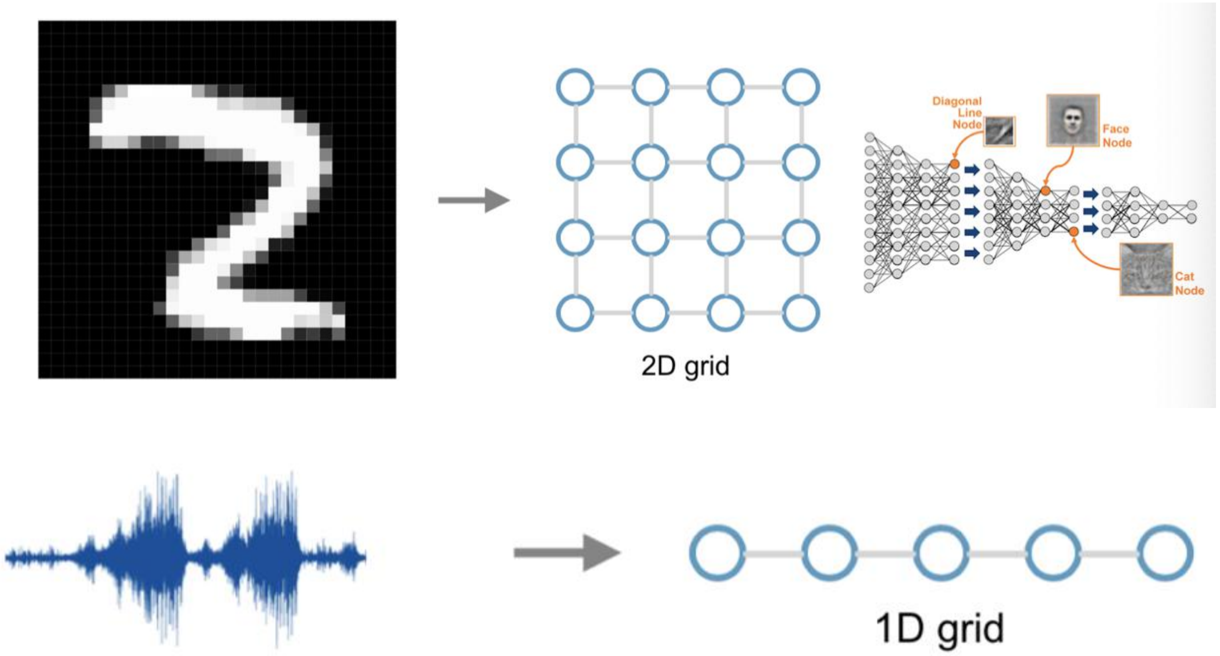
三角不等式

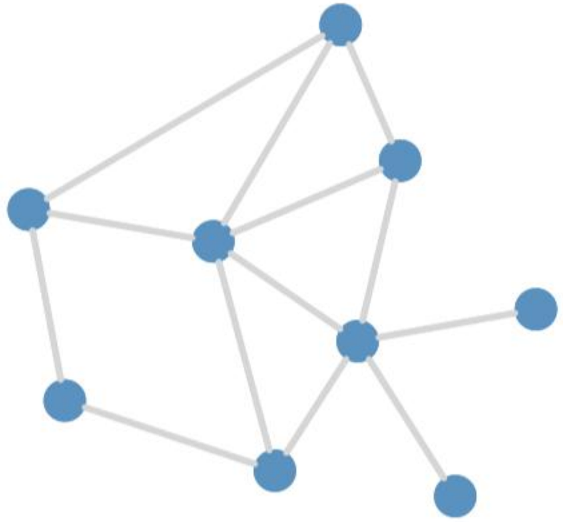
图表示学习

图表示学习的难点

欧式空间的网格数据

非欧式空间的图数据





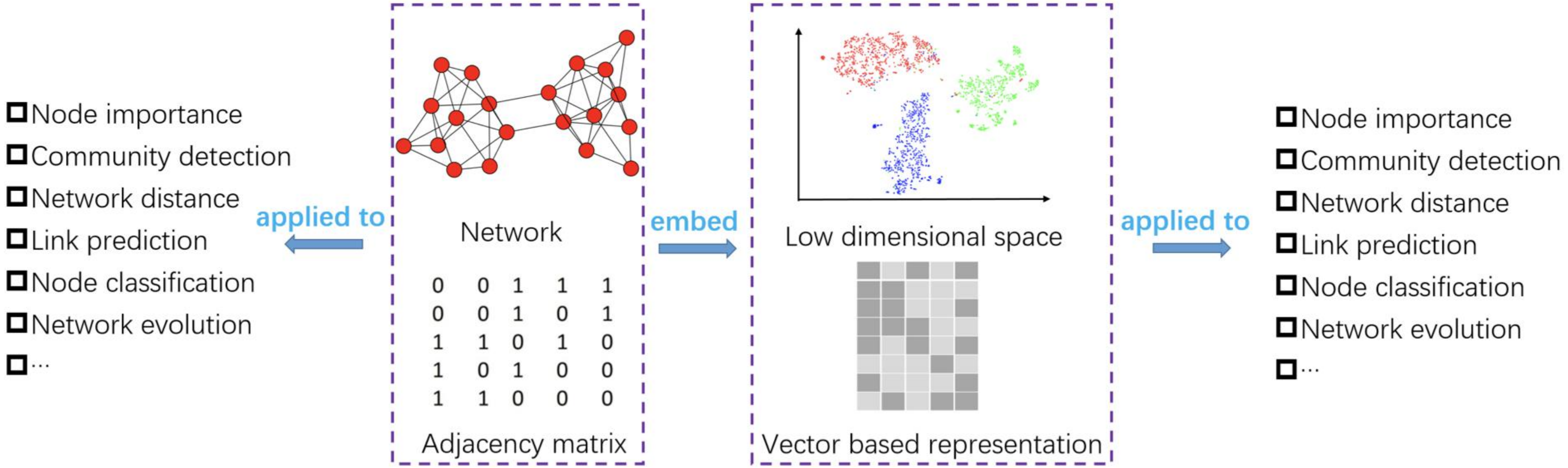
结构固定且简单

结构灵活且复杂

图表示学习

图表示学习的定义:

图表示学习: 给定一个图$\mathcal{G}=(\mathcal{V},\mathcal{E})$,图表示学习旨在基于一个参数化模型来学习$d$维的节点表示$e_{\mathcal{V}}\in\mathbb{R}^{|V|\times d}(d\ll|V|)$, 该$d$维表示能够保持节点的结构和性质,并为下游任务提供良好的特征表示.



图表示学习

异步社区
人民邮电出版社

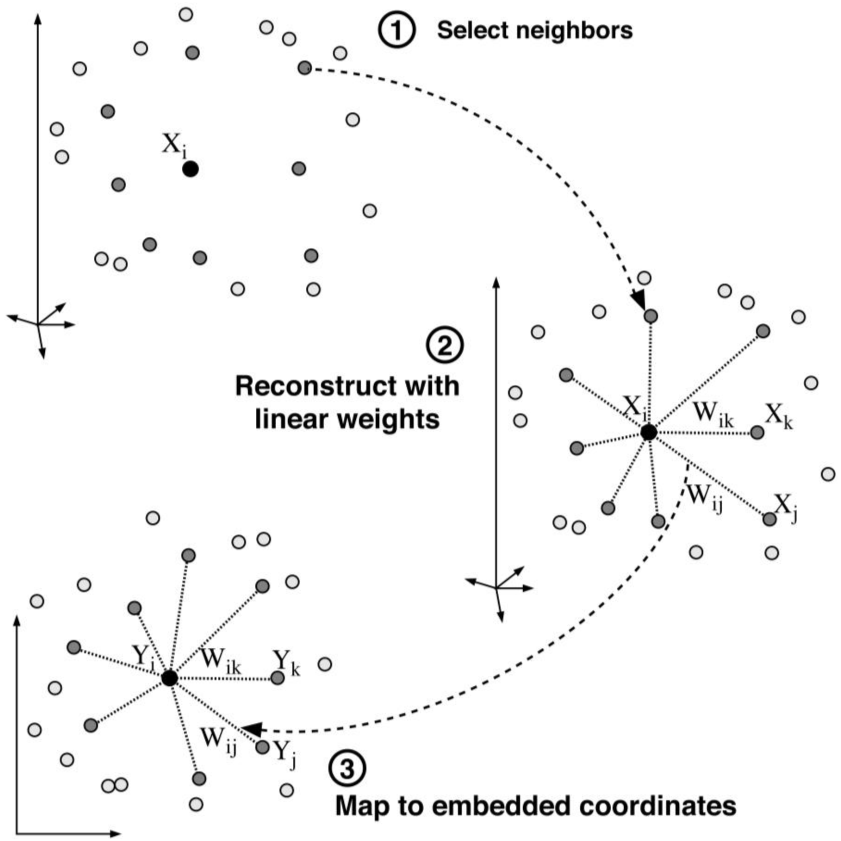
图表示学习分类

- 1.0阶段: 基于拓扑结构的图表示学习
  - LLE, Laplacian Eigenmaps
- 2.0阶段: Graph Embedding
  - 分解模型: GraphRep等
  - 随机游走模型: DeepWalk, node2vec
  - 深层模型: SDNE
- 3.0阶段: 图神经网络 Graph Neural Network
  - 图卷积神经网络 Graph Convolutional Network
  - 图注意力网络 Graph Attention Network

图表示学习

图表示学习分类

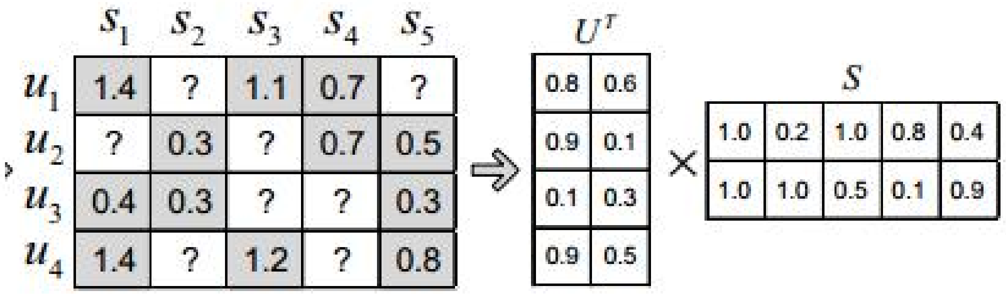
- 1.0阶段: 基于拓扑结构的图表示学习
  - LLE, Laplacian Eigenmaps

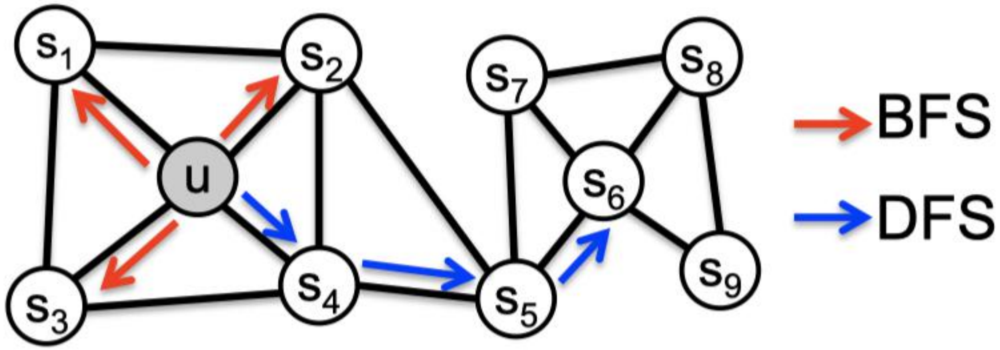


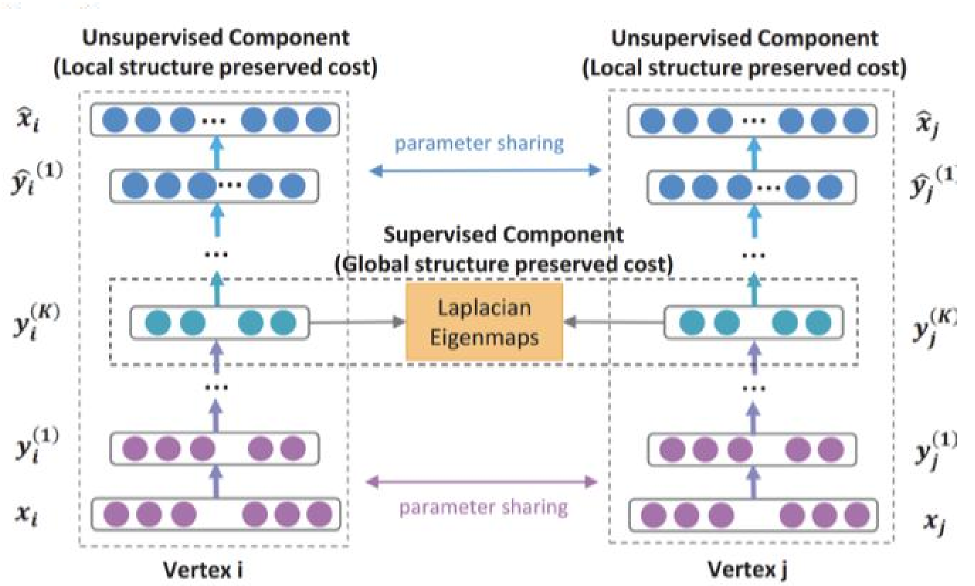
图表示学习

图表示学习分类

- 2.0阶段: Graph Embedding
  - 分解模型: GraphRep等
  - 随机游走模型: DeepWalk, node2vec
  - 深层模型: SDNE



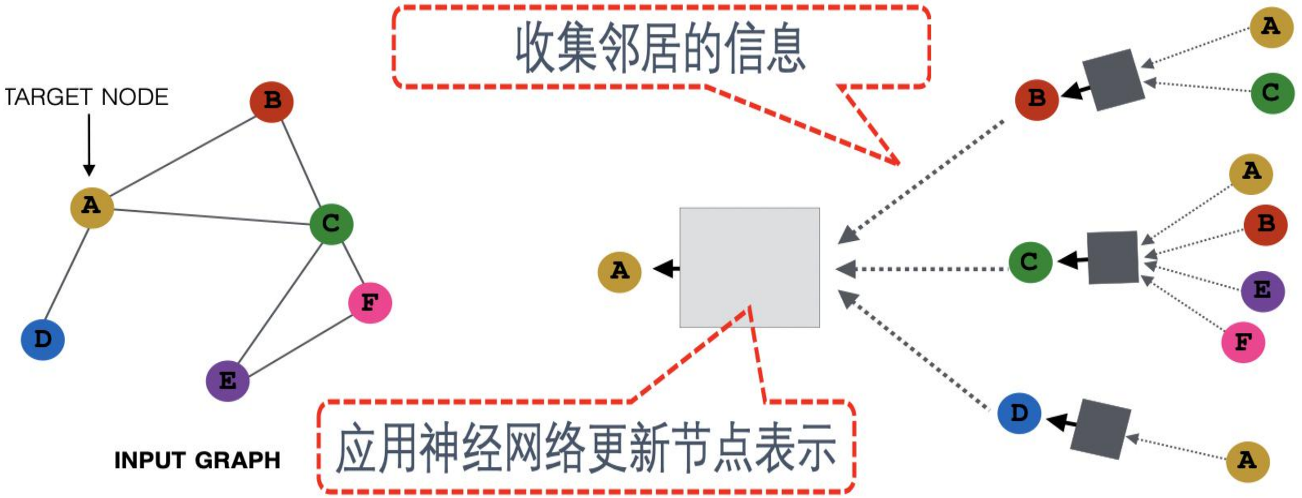


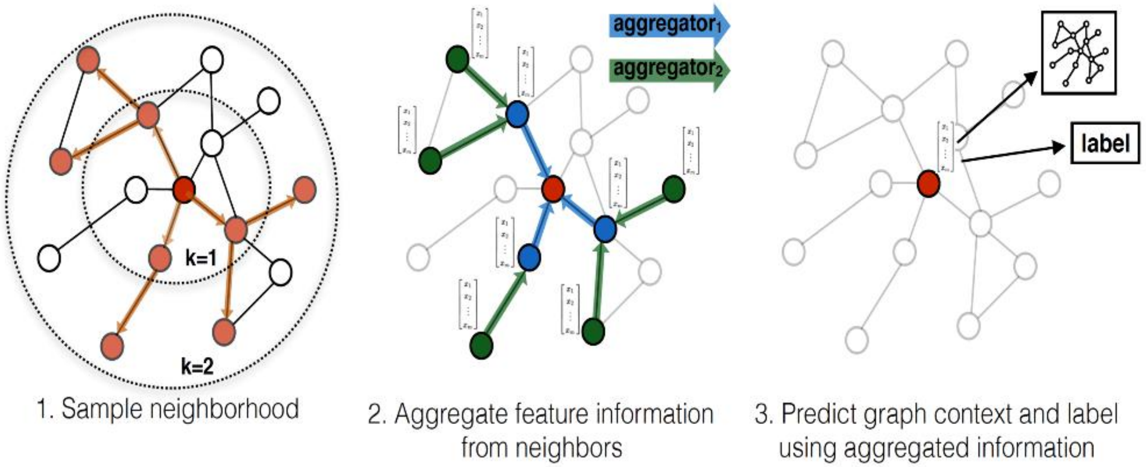


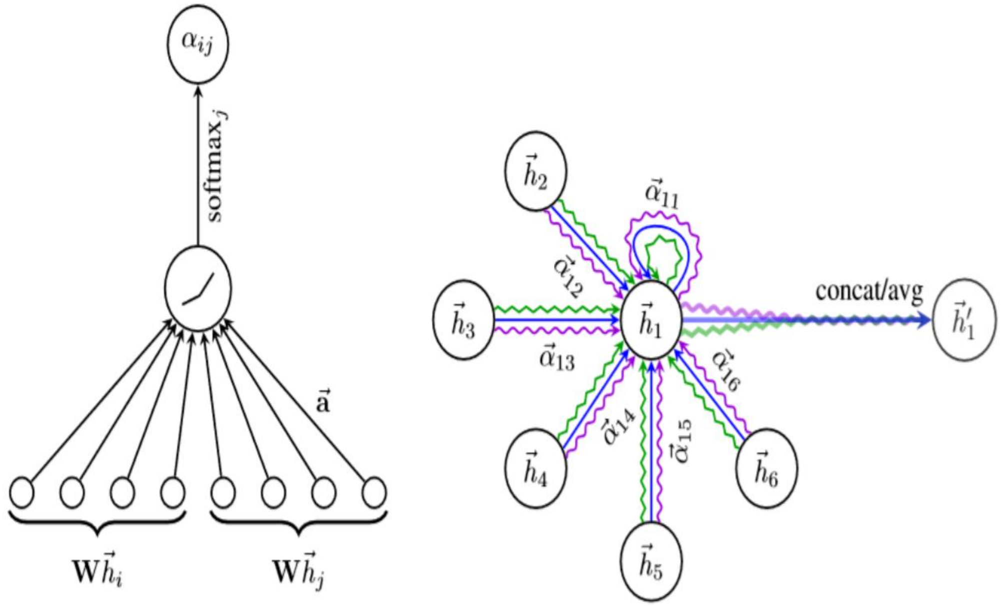
图表示学习

图表示学习分类

- 3.0阶段: 图神经网络
  - 图卷积网络Graph Convolutional Network
  - 图注意力网络Graph Attention Network



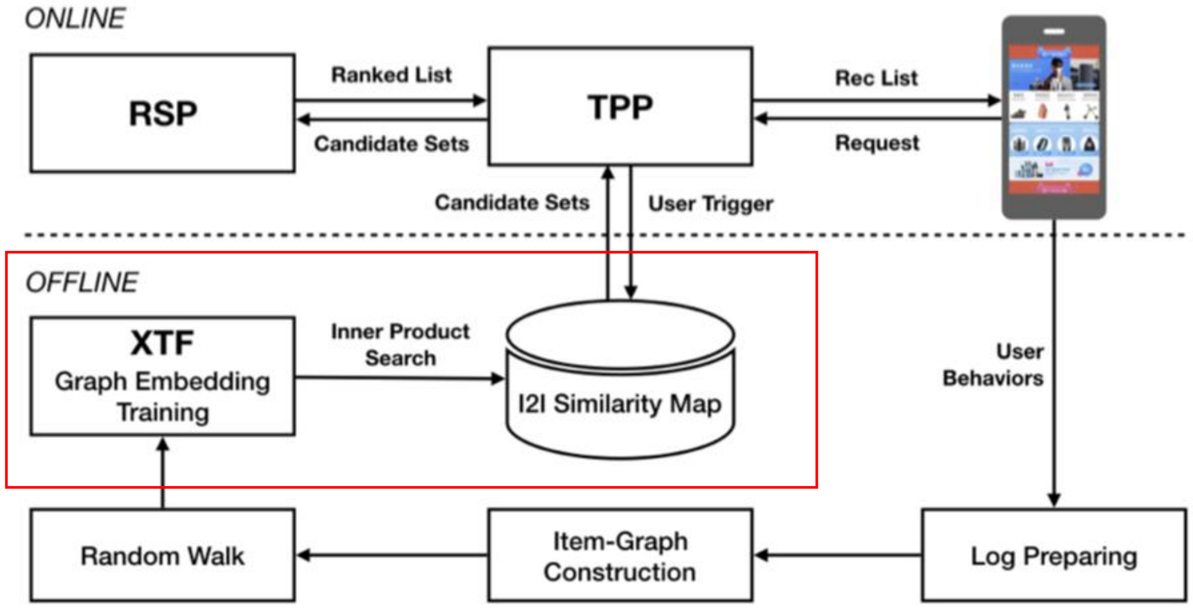




图表示学习

图表示学习在推荐系统中的应用(用户/商品表示, 物品召回, 个性化推荐)

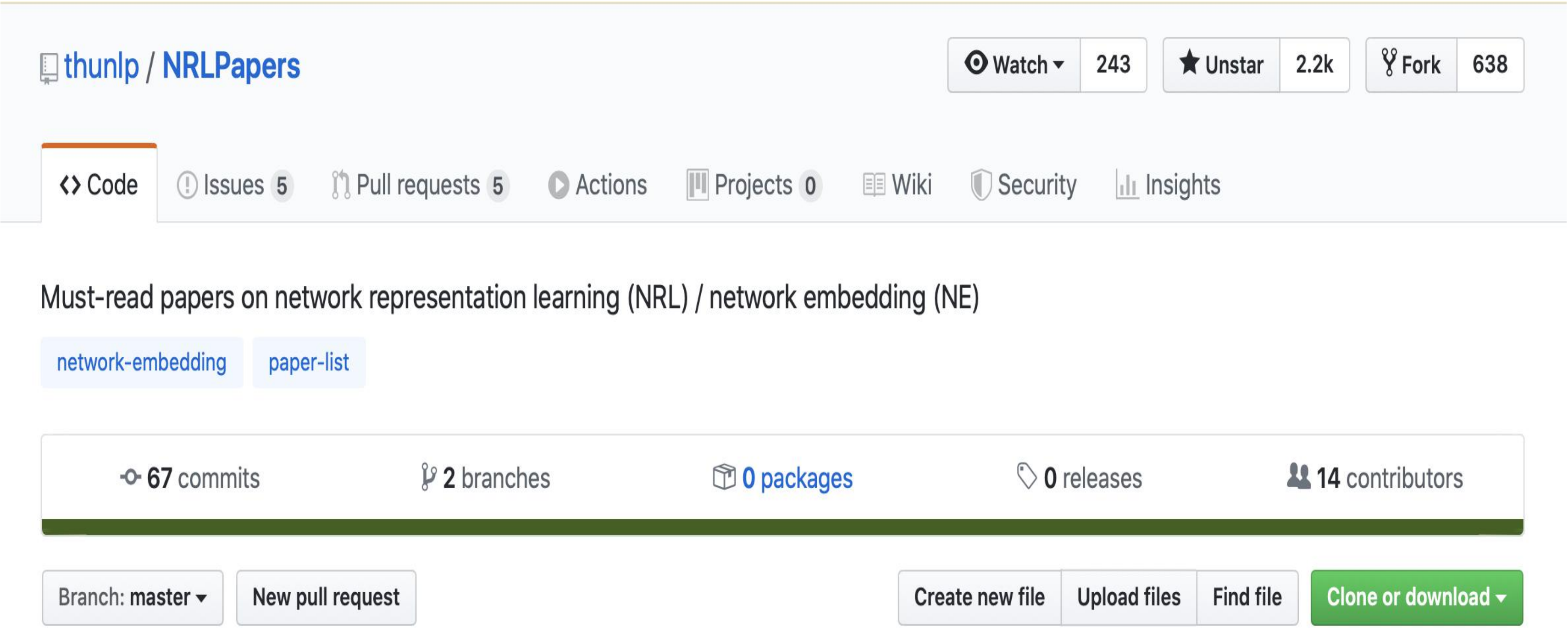
- KDD 2018 Best Paper Real-time Personalization using **Embeddings** for Search Ranking at Airbnb
- KDD 2018 Billion-scale Commodity **Embedding** for E-commerce Recommendation in Alibaba



图表示学习

相关资源

https://github.com/thunlp/NRLPapers



Content

1. Survey Papers
2. Models
   i. Bacis Models
   ii. Attributed Network
   iii. Dynamic Network
   iv. Heterogeneous Information Network
   v. Bipartite Network
   vi. Directed Network
   vii. Other Models
3. Applications
   i. Natural Language Processing
   ii. Knowledge Graph
   iii. Social Network
   iv. Graph Clustering
   v. Community Detection
   vi. Recommendation
   vii. Other Applications

02 图表示学习 1.0

图表示学习分类 1.0:

- LLE （Science 2000）
- Laplacian Eigenmaps （NIPS 2001）
- Cauchy Graph Embedding （ICML 2011）

Laplacian Eigenmaps and Spectral Techniques for Embedding and Clustering

Mikhail Belkin and Partha Niyogi
Depts. of Mathematics and Computer Science
The University of Chicago
Hyde Park, Chicago, IL 60637.
(misha@math.uchicago.edu,niyogi@cs.uchicago.edu)

Nonlinear Dimensionality Reduction by Locally Linear Embedding

Sam T. Roweis[1] and Lawrence K. Saul[2]

Many areas of science depend on exploratory data analysis and visualization. The need to analyze large amounts of multivariate data raises the fundamental problem of dimensionality reduction: how to discover compact representations of high-dimensional data. Here, we introduce locally linear embedding (LLE), an unsupervised learning algorithm that computes low-dimensional, neighborhood-preserving embeddings of high-dimensional inputs. Unlike clustering methods for local dimensionality reduction, LLE maps its inputs into a single global coordinate system of lower dimensionality, and its optimizations do not involve local minima. By exploiting the local symmetries of linear reconstructions, LLE is able to learn the global structure of nonlinear manifolds, such as those generated by images of faces or documents of text.

Cauchy Graph Embedding

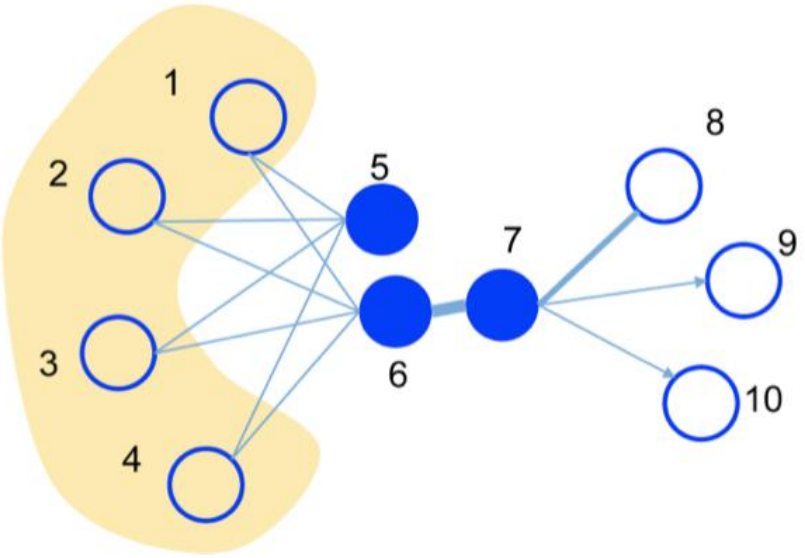
Dijun Luo DIJUN.LUO@GMAIL.COM
Chris Ding CHQDING@UTA.EDU
Feiping Nie FEIPINGNIE@GMAIL.COM
Heng Huang HENG@UTA.EDU
The University of Texas at Arlington, 701 S. Nedderman Drive, Arlington, TX 76019

图表示学习

图表示学习分类 1.0:

- LLE （Science 2000）
- Laplacian Eigenmaps （NIPS 2001）
- Cauchy Graph Embedding （ICML 2011）

上述算法均主要保持图上的**一阶结构**。换句话说，如果两个节点有边直接相连，那么这两个节点的向量表示也会非常像。



图表示学习: 给定一个图$\mathcal{G} = (\mathcal{V}, \mathcal{E})$,图表示学习旨在基于一个参数化模型来学习$d$维的节点表示$e_{\mathcal{V}} \in \mathbb{R}^{|V| \times d}(d \ll |V|)$, 该$d$维表示能够保持节点的结构和性质,并为下游任务提供良好的特征表示.

Locally Linear Embedding (LLE) (Science 2000)

一句话描述LLE: 每个节点都是由其邻居的加权平均来表示.

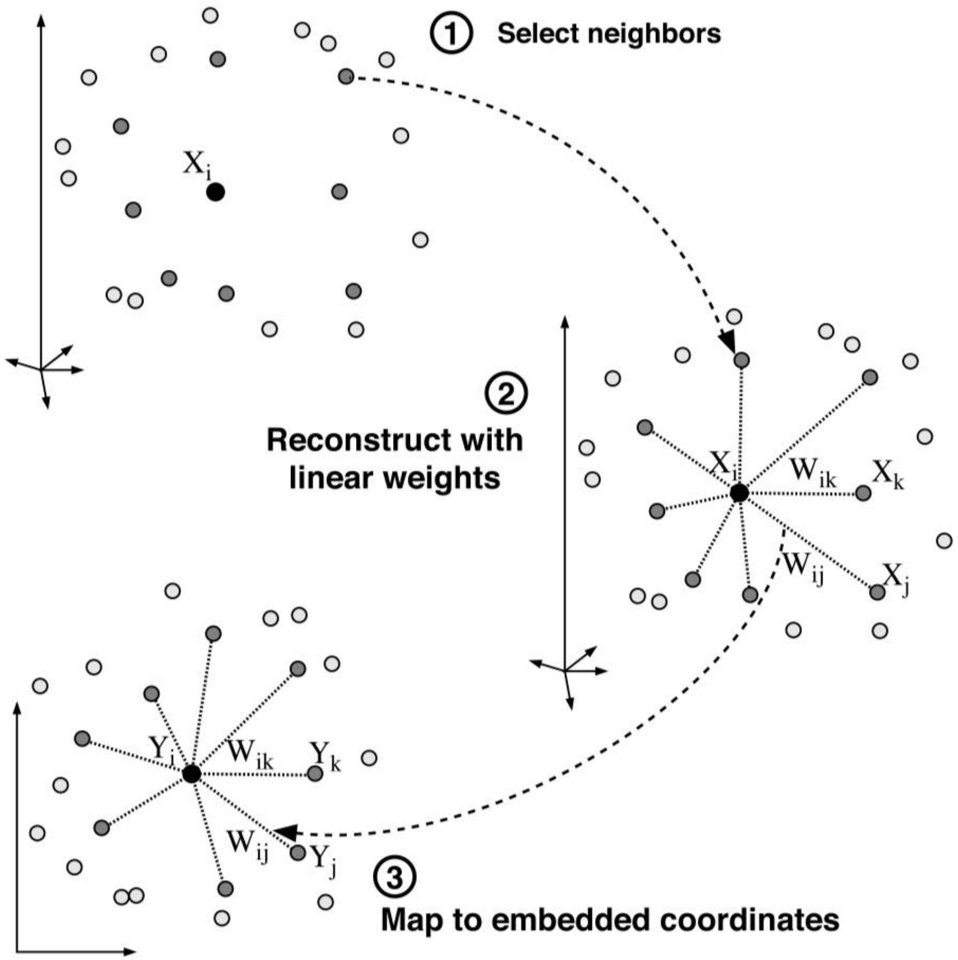
- 节点i与其邻居j建立了联系(有边相连 $W_{ij} > 0$)，则存在某种相似性。
  - 边权重越大， 节点对的相似性越高。
- 每个邻居j对于节点i的重要性W_{ij}，应该有所不同。
- 这里节点与邻居都用d维向量来表示，分别为Y_i和Y_j。

$$Y_i \approx \sum_j W_{ij} Y_j \quad \forall i \in V \, , \sum_j W_{ij} = 1$$

- 优化目标

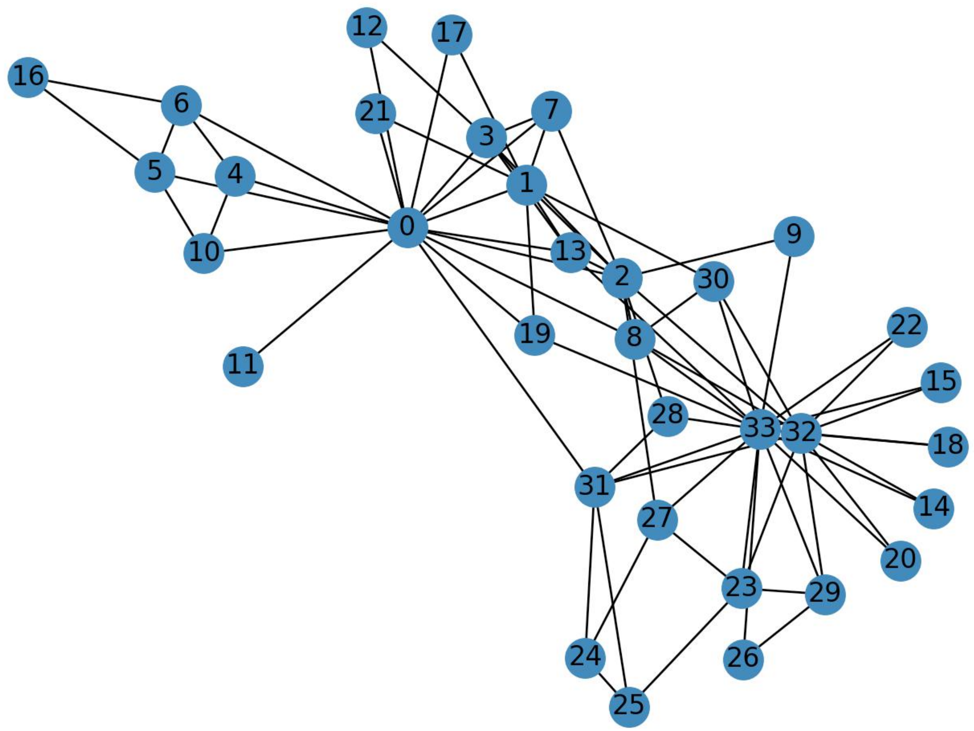
$$\min \phi(Y) = \sum_i |Y_i - \sum_j W_{ij} Y_j|^2$$

- 复杂度为$O\left(|E|d^2\right)$，和边的数量成线性。
- 主要保持了图上的一阶结构(直接相连的邻居关系)。
- 对于原始形态不是图结构的数据， 可以通过KNN等方式建图，然后套用LLE。

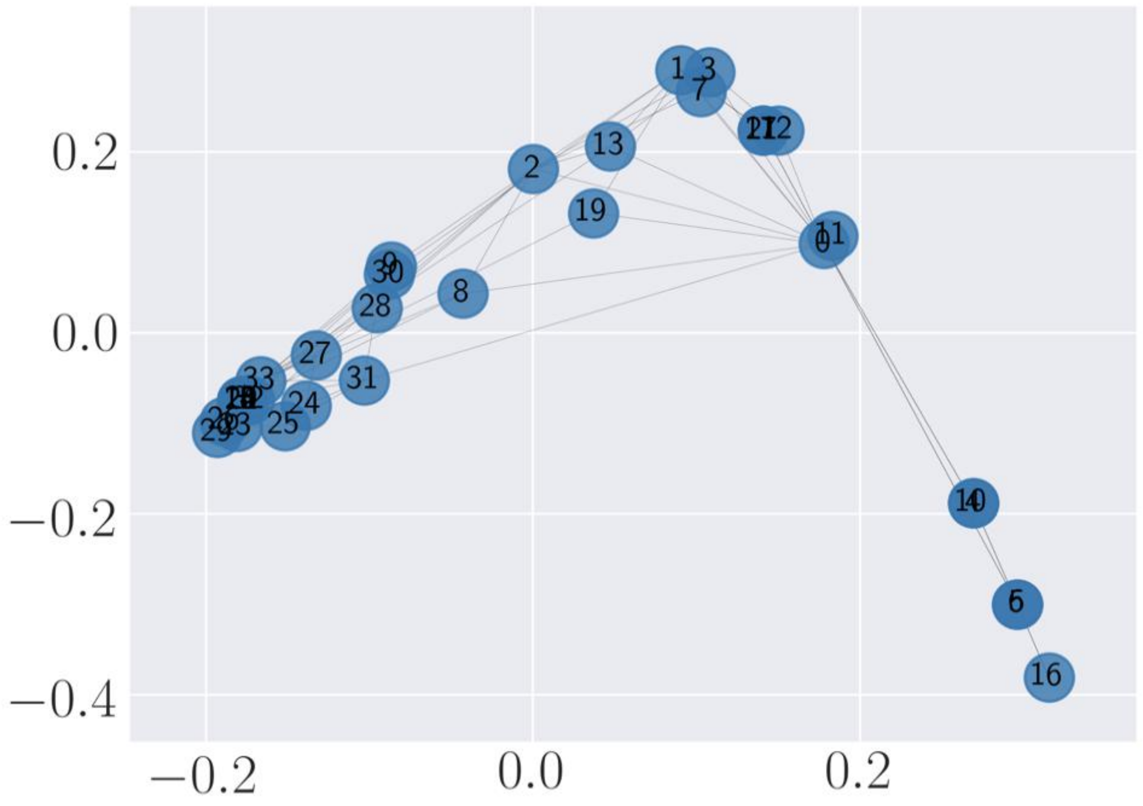


Locally Linear Embedding (LLE) (Science 2000)

Visualization of Karate graph using Locally Linear Embedding



34位空手道运动员构成的图



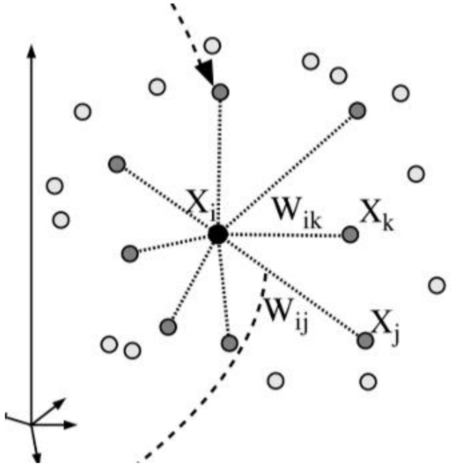
34位空手道运动员的2d表示

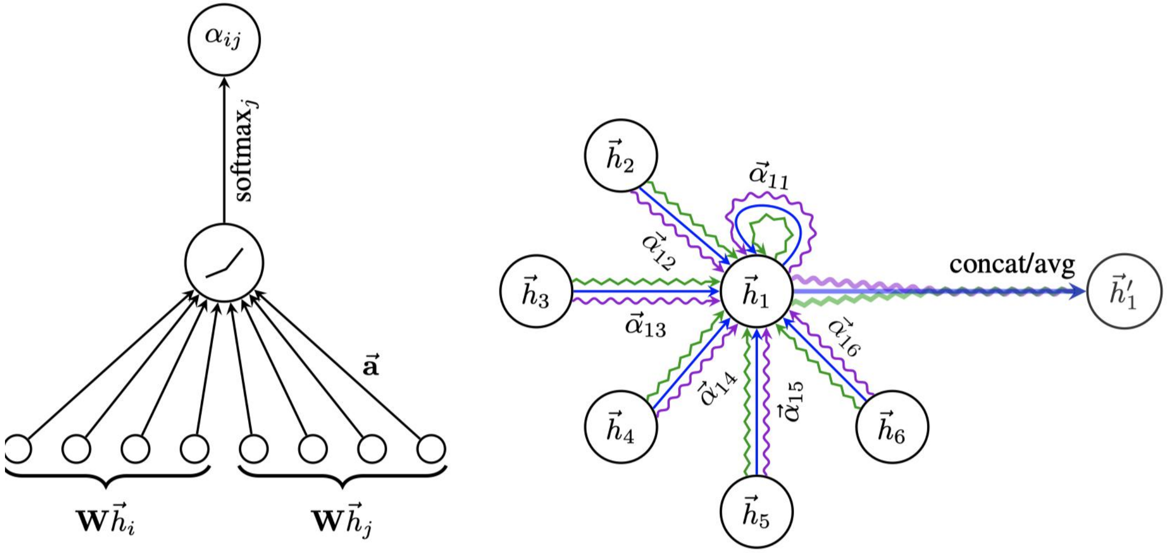
图表示学习

Locally Linear Embedding (LLE) (Science 2000)

思考：

LLE能不能更进一步？利用神经网络+注意力机制来学习节点i与其邻居j的权重 W_{ij}呢？





$$Y_i \approx \sum_j W_{ij} Y_j \quad \forall i \in V \, , \sum_j W_{ij} = 1$$

$$H_i = \sigma\left(\sum_{j \in \mathcal{N}_i} \alpha_{ij} H_j\right) , \alpha_{ij} = Attention(H_i \, , H_j)$$

2000Science Locally Linear Embedding

2018ICLR Graph Attention Network

图表示学习

Laplacian Eigenmaps (NIPS 2001)

一句话：节点i与其邻居j之间边的权重 W_{ij}越大， 则它们的表示Y_i和Y_j越像（距离越近）。

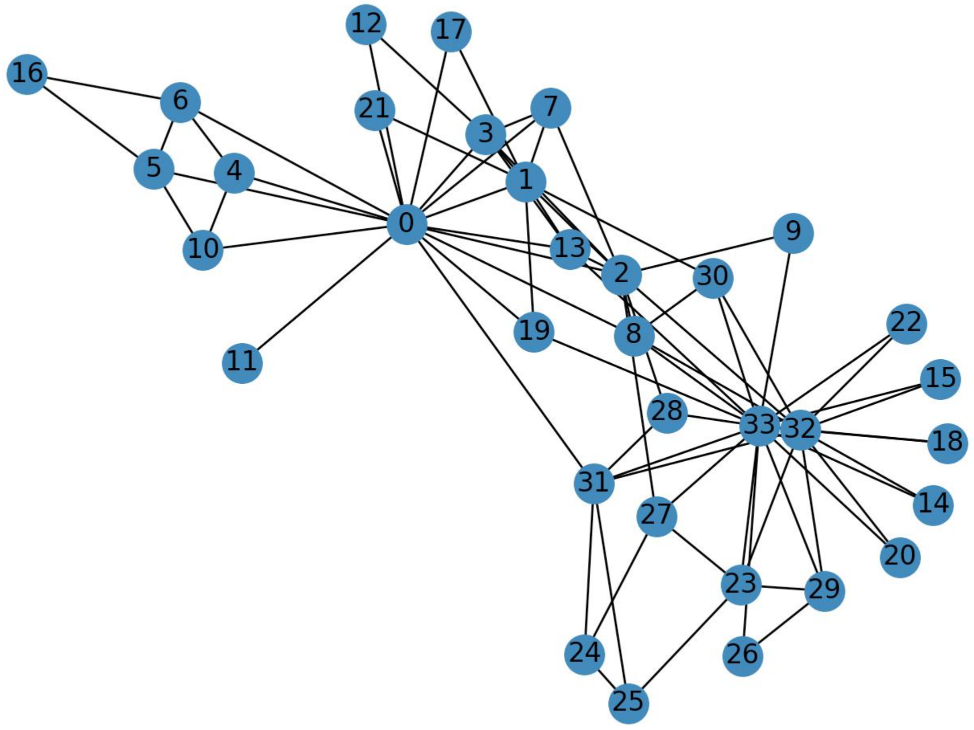
$$\min \phi(Y)=\frac{1}{2}\sum_{i,j}|Y_i-Y_j|^2 W_{ij}$$

- $|Y_i-Y_j|^2$代表了节点间的距离，是一个二次方惩罚函数。
- 同样距离（固定$|Y_i-Y_j|^2$），权重W_{ij}越大，惩罚函数$\phi(Y)$越大。
- 复杂度为$O\left(|E|d^2\right)$，和边的数量成线性。
- 主要保持了图上的一阶结构(直接相连的邻居关系)。

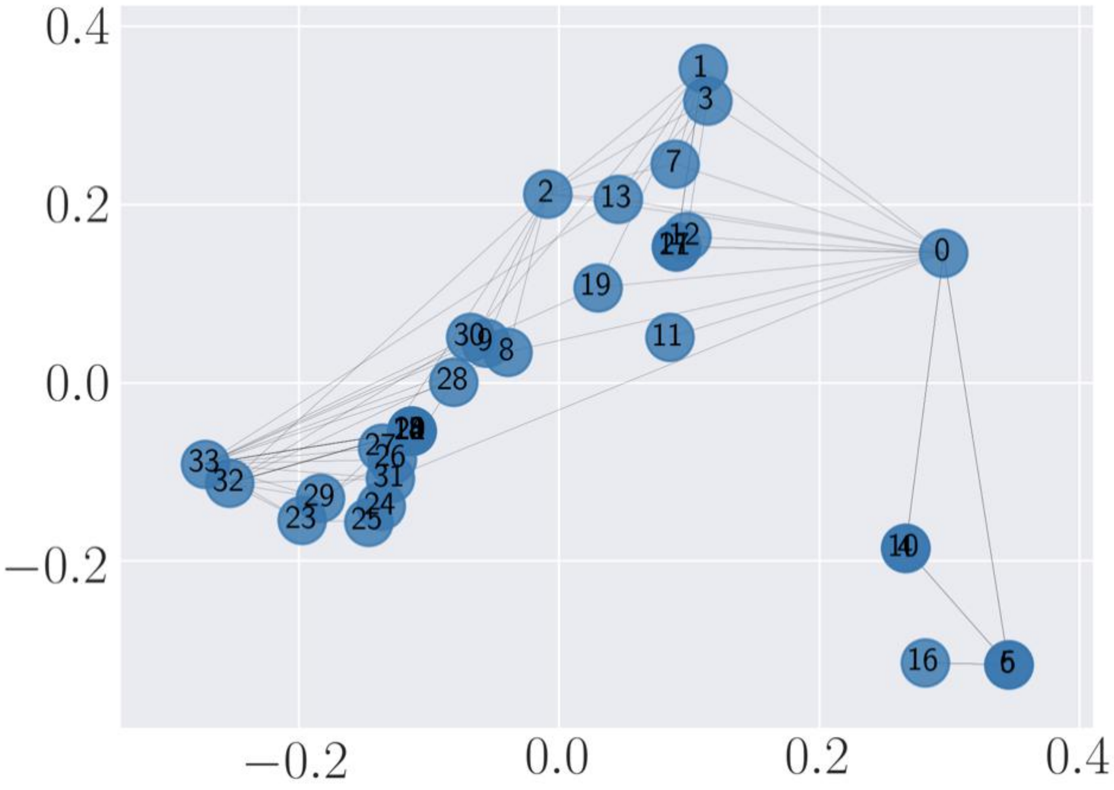
图表示学习

Laplacian Eigenmaps (NIPS 2001)

Visualization of Karate graph using Laplacian Eigenmaps



34位空手道运动员构成的图



34位空手道运动员的2d表示

图表示学习

Cauchy Graph Embedding （ICML 2011）

与Laplacian Eigenmaps类似， Cauchy Graph Embedding也是：节点i与其邻居j之间边的权重 W_{ij}越大，则它们的表示Y_i和Y_j越像（距离越近）。

$$\max \phi(Y) = \sum_{i,j} \frac{W_{ij}}{|Y_i - Y_j|^2 + \sigma^2}$$

$|Y_i - Y_j|^2$代表了节点间的距离，是一个二次方惩罚函数，**但是放在分母的位置。**

- 复杂度为$O\left(|E|d^2\right)$，和边的数量成线性。
- 主要保持了图上的一阶结构(直接相连的邻居关系)。

为什么叫做Cauchy Graph Embedding?

- 因为$f(x) = 1/\left(x^2 + \sigma^2\right)$是Cauchy distribution.

节点对之间的距离函数有多种选择：

- 二次方 $|Y_i - Y_j|^2$
- 绝对值 $|Y_i - Y_j|$
- 高斯 $\exp\left[-\frac{(Y_i - Y_j)^2}{\sigma^2}\right]$
- 指数 $\exp\left[-\frac{(Y_i - Y_j)}{\sigma}\right]$