Graph Embedding for Social Network Analysis

Jiani Zhang & Irwin King

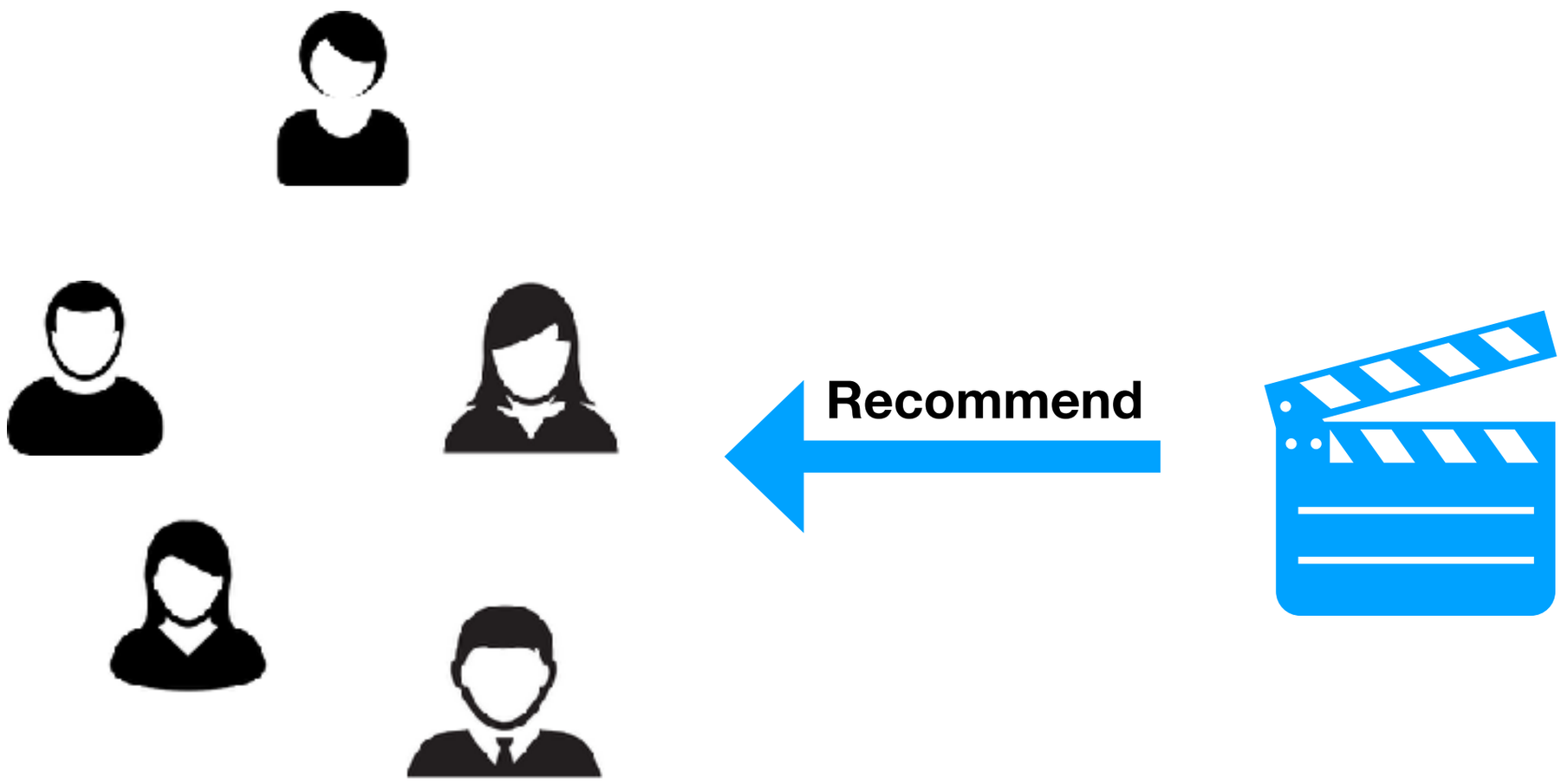
Director, Centre for eLearning Innovation and Technology (ELITE), CUHK
Director, Shenzhen Rich Media and Big Data Analytics and Application Key Lab, SZRI, CUHK
PI, The Knowledge and Education Exchange Platform (KEEP), VeriGuide System, CUHK
Former Associate Dean (Education), Faculty of Engineering, CUHK

Department of Computer Science and Engineering
The Chinese University of Hong Kong

king@cse.cuhk.edu.hk
http://www.cse.cuhk.edu.hk/~king

电影推荐

Movie Recommendation

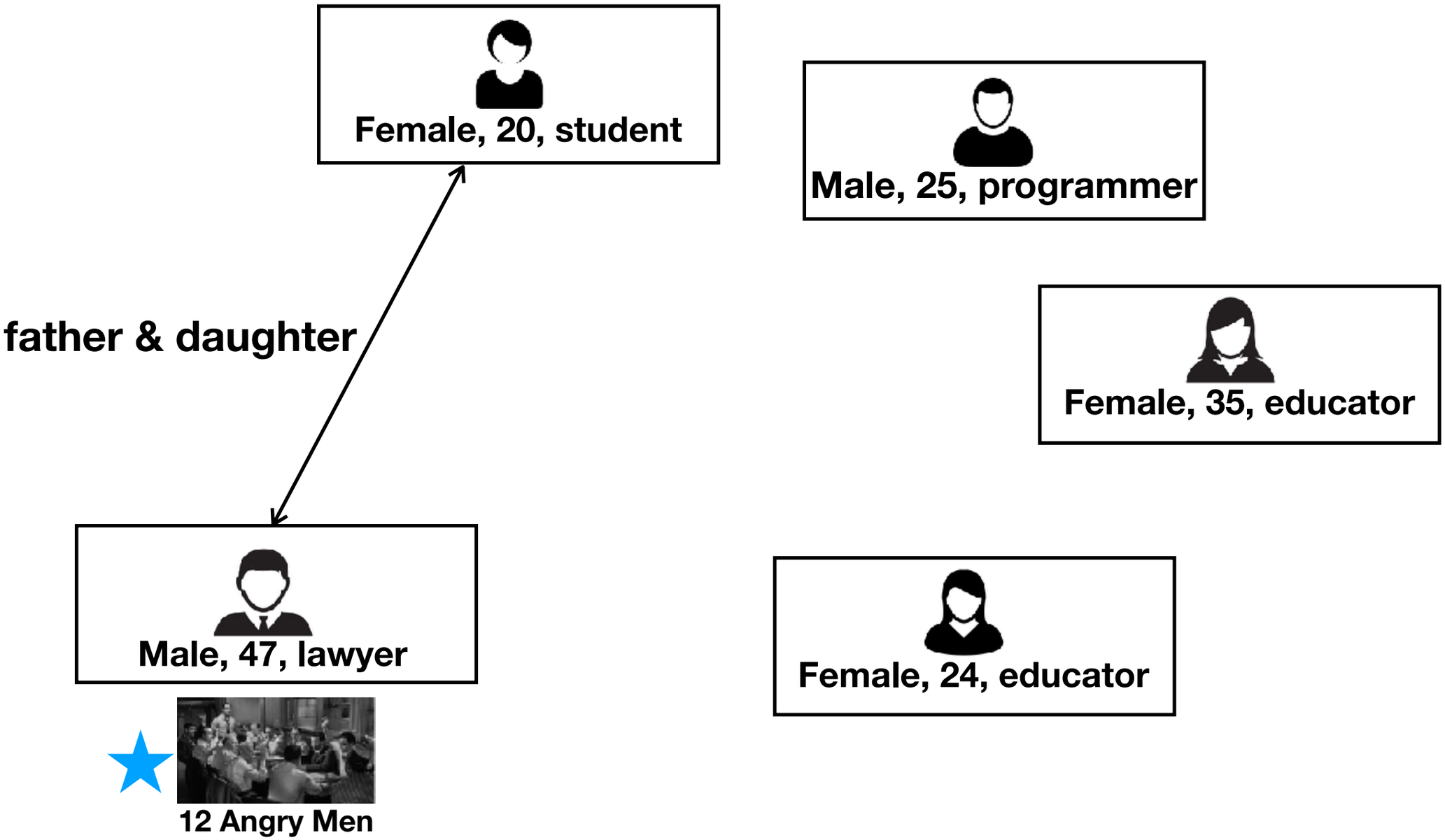


电影推荐

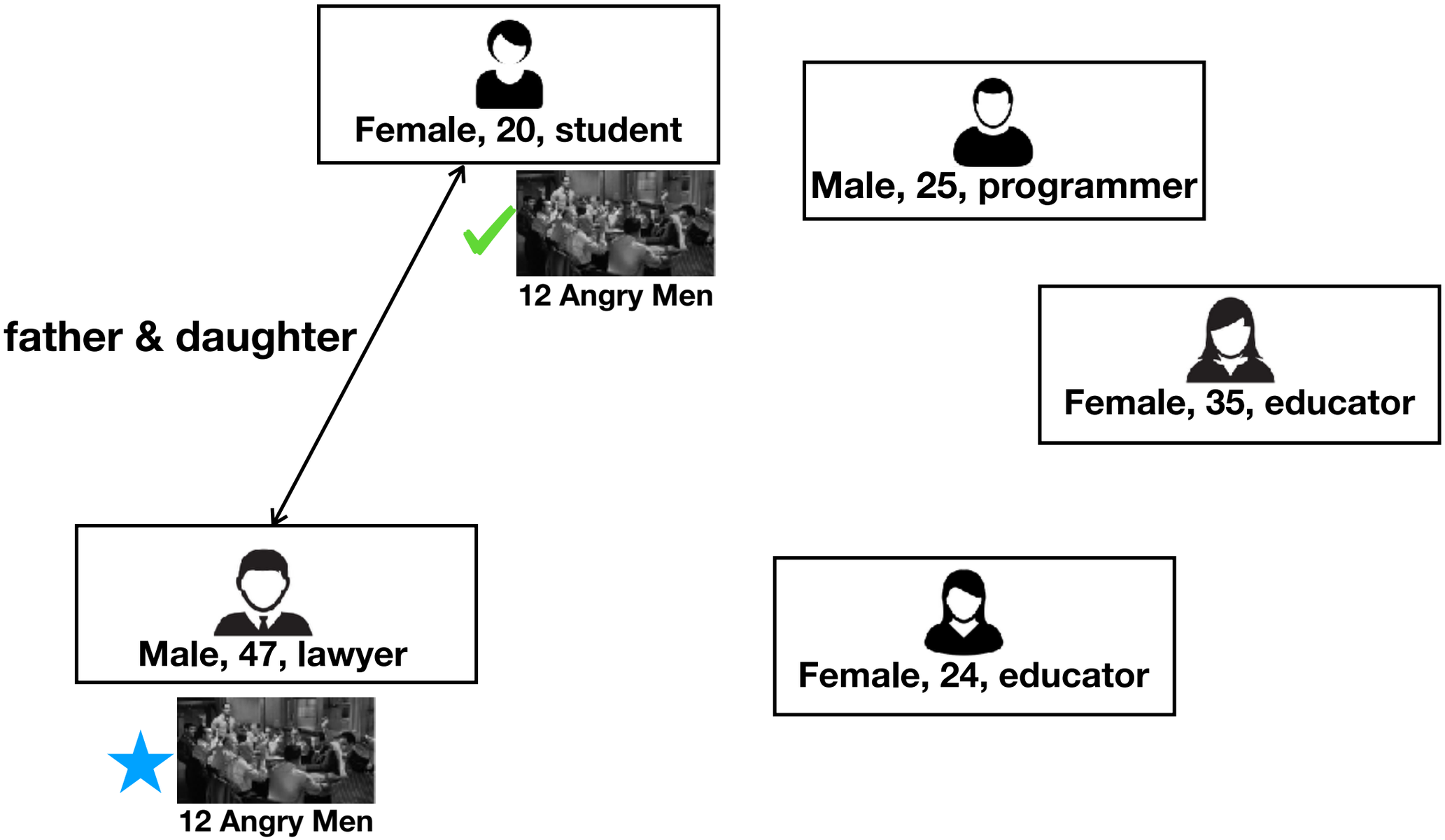
Movie Recommendation

	 12 Angry Men	 Star War VII	 Star War VIII	 Dead Poets Society	 The Princess Diaries
Female, 20, student					★
Male, 25, programmer		★			
Female, 35, educator				★	
Female, 24, educator					★
Male, 47, lawyer	★				

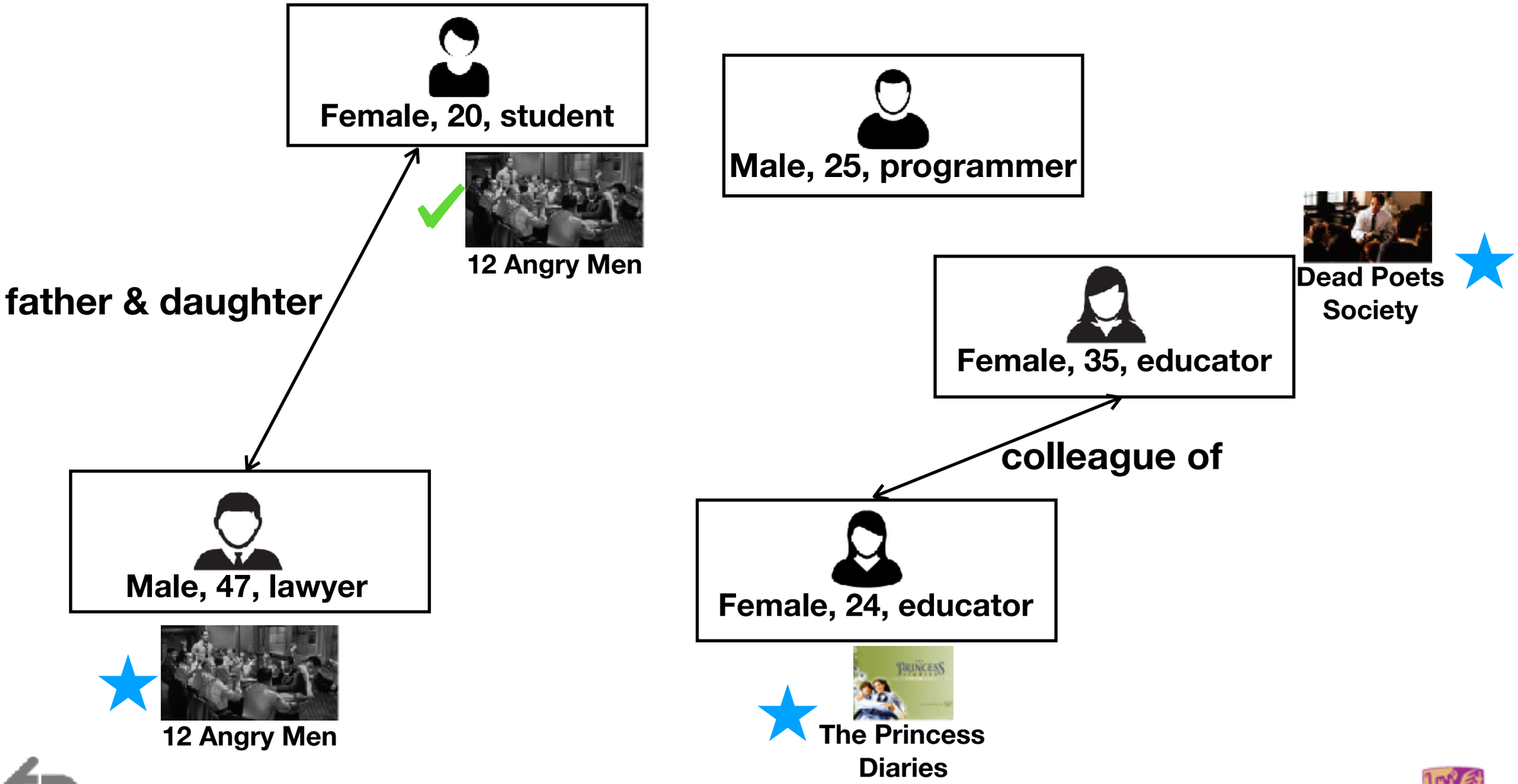
引入社交网络
With Social Network



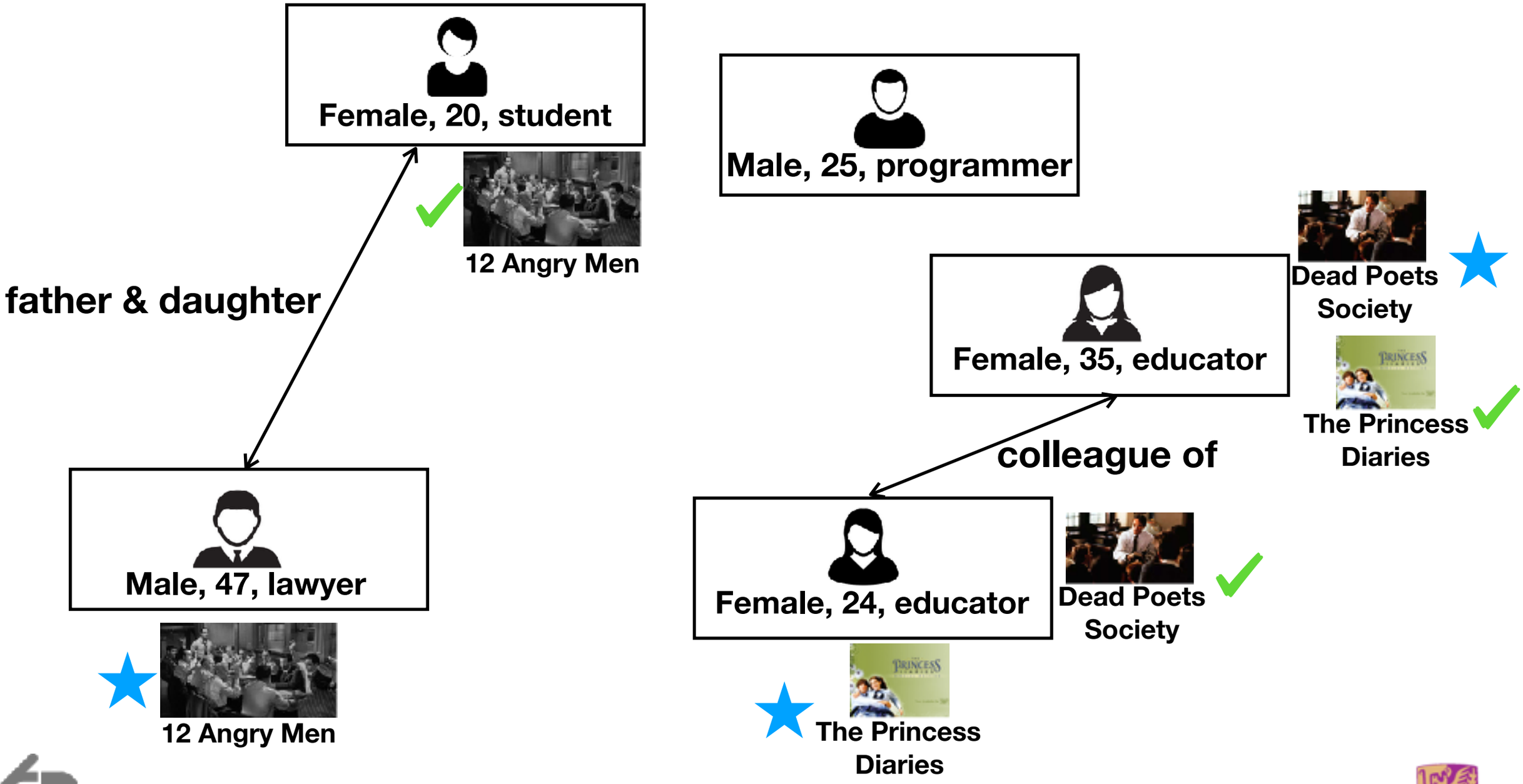
引入社交网络
With Social Network



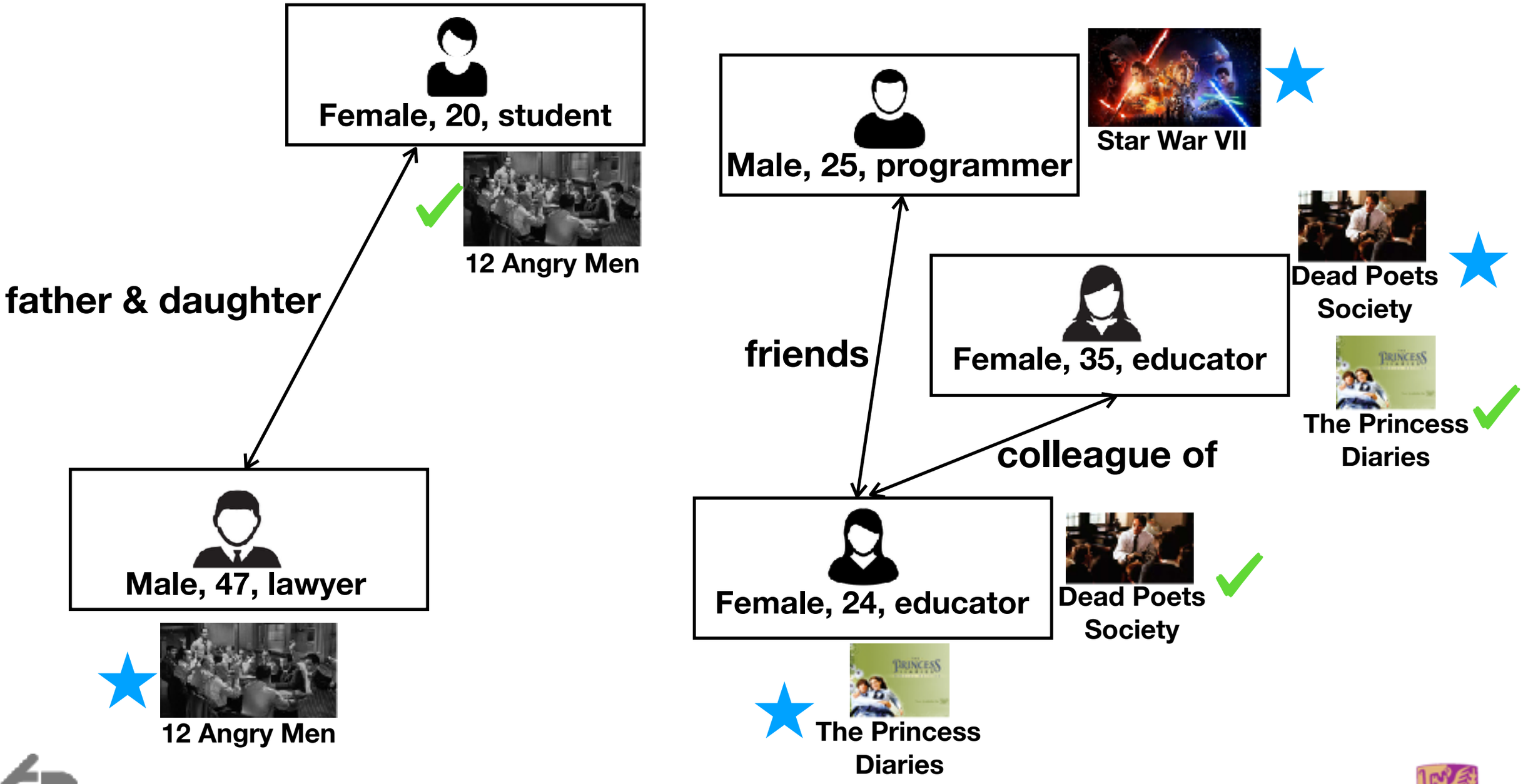
引入社交网络
With Social Network



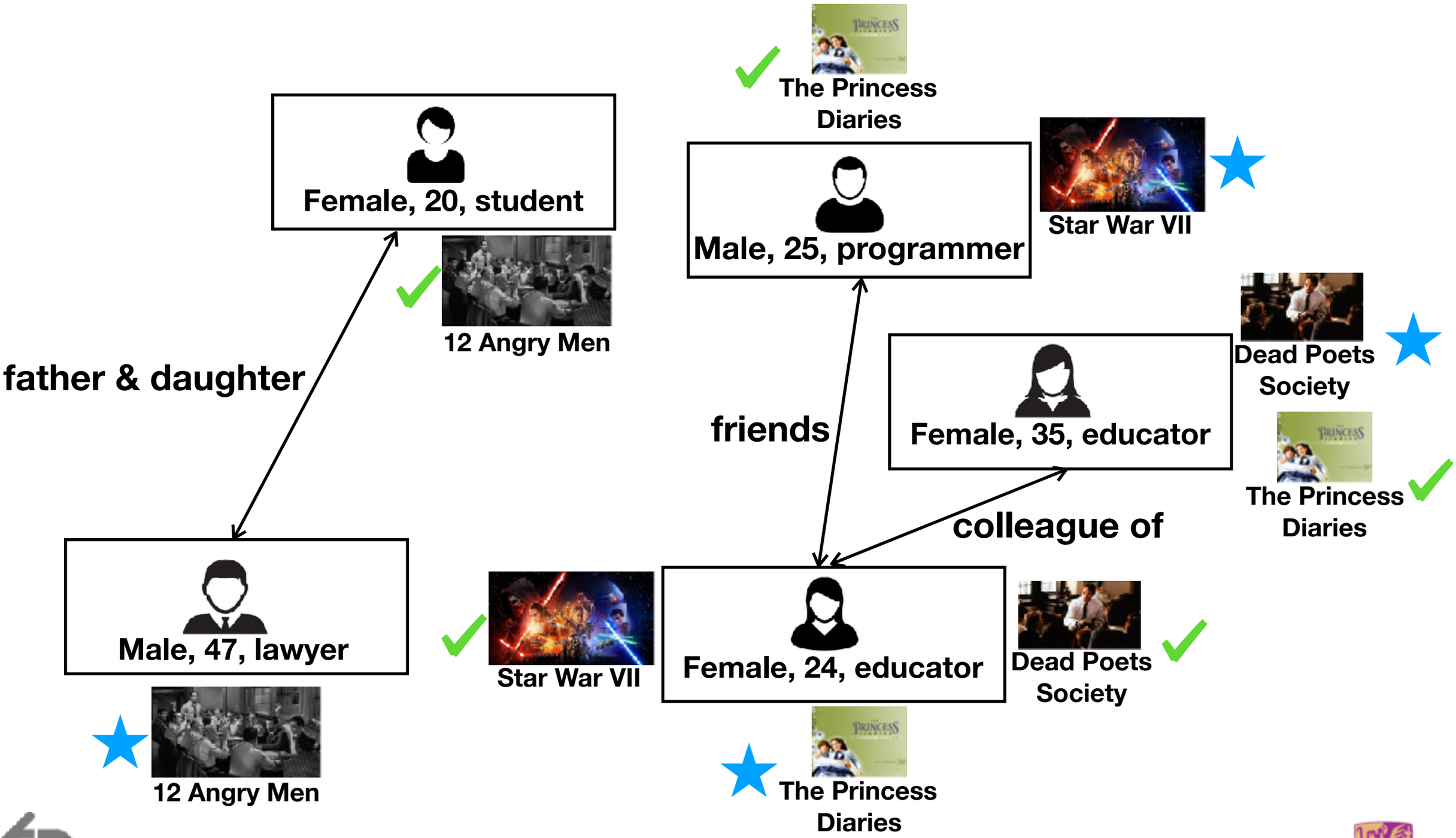
引入社交网络
With Social Network



引入社交网络
With Social Network



引入社交网络
With Social Network



电影推荐

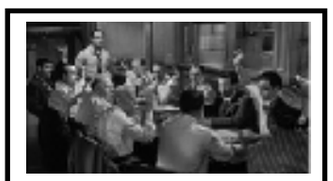
Movie Recommendation

	 12 Angry Men	 Star War VII	 Star War VIII	 Dead Poets Society	 The Princess Diaries
 Female, 20, student	✓				★
 Male, 24, programmer		★			✓
 Female, 35, educator				★	✓
 Female, 24, educator		✓		✓	★
 Male, 47, lawyer	★				

引入内容网络
With Content Network


Star War VII

The same series


12 Angry Men

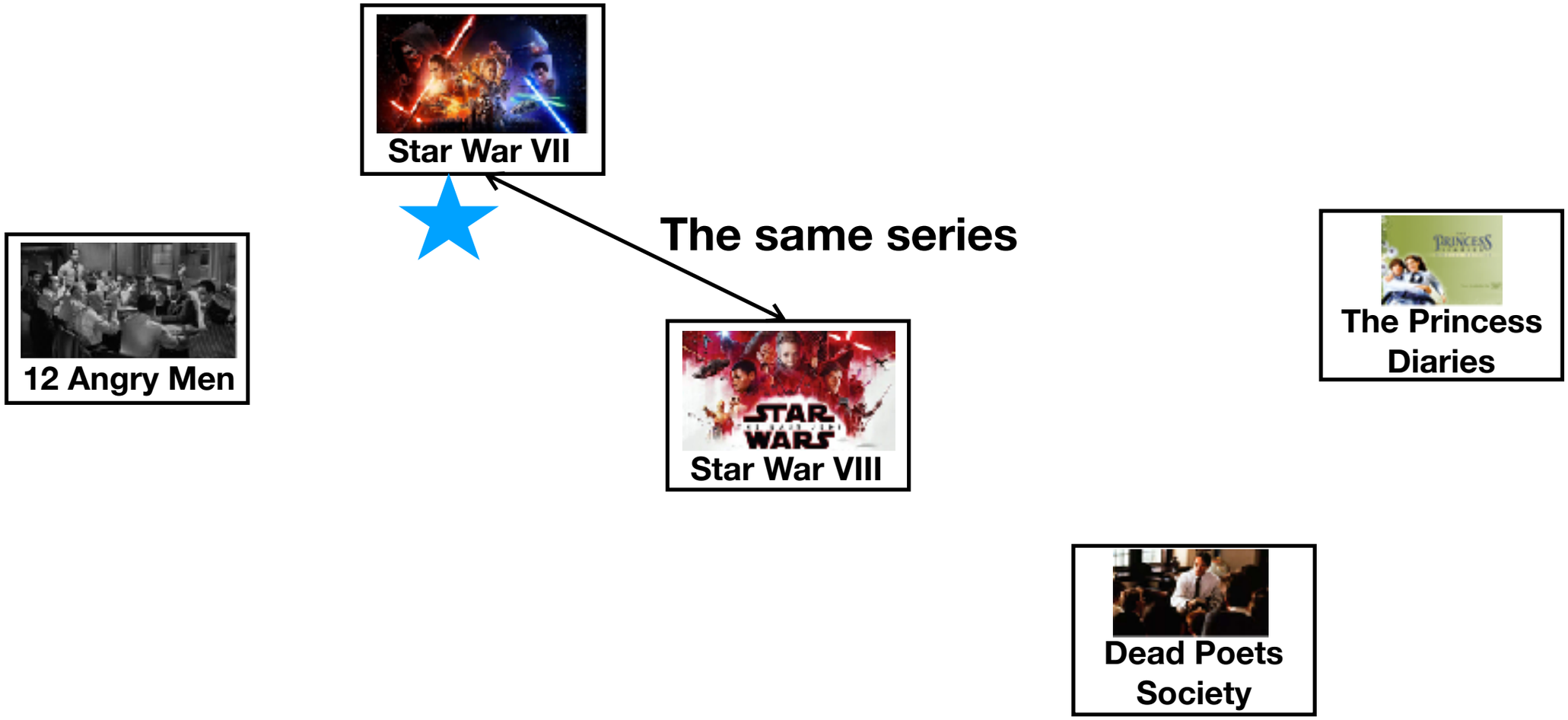

Star War VIII


The Princess Diaries


Dead Poets Society

引入内容网络

With Content Network



引入内容网络

With Content Network

Star War VII

关系：同一系列电影

The same series

12 Angry Men

Star War VIII

The Princess Diaries

Dead Poets Society

电影推荐

Movie Recommendation

	 12 Angry Men	 Star War VII	 Star War VIII	 Dead Poets Society	 The Princess Diaries
 Female, 20, student	✓				★
 Male, 24, programmer		★	✓		✓
 Female, 35, educator				★	✓
 Female, 24, educator		✓		✓	★
 Male, 47, lawyer	★				

社交网络

Social Networks

社交网络天然就是以graph(图)形式存在

- Social Networks are intrinsically in the form of **graphs**
  - **Nodes** (individuals) and ties **节点：人或者其他连接介质**
  - **Edges** or **Links** (relationships or interactions) **边：节点关系或者交互**

相关应用 Some Useful Tasks

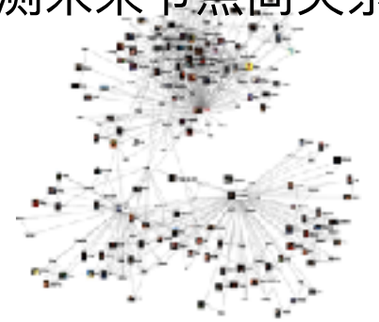
节点分类
Node Classification

- To predict the role function of a node

- 分析节点的分类、功能等。

关系预测
Link Prediction

- To predict future possible links in the network
- 预测未来节点间关系的变化



推荐
Recommendation

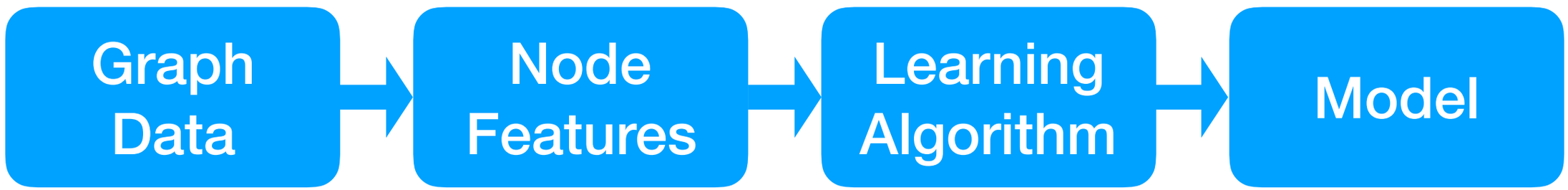
- To analyze users' preference patterns to suggest potential

社区检测
Community Detection

- To encode a partition of nodes into multiple communities
- 将不同节点划分到不同社区

图学习流程 Graph Learning Lifecycle

图数据 节点特征 学习算法 最终模型



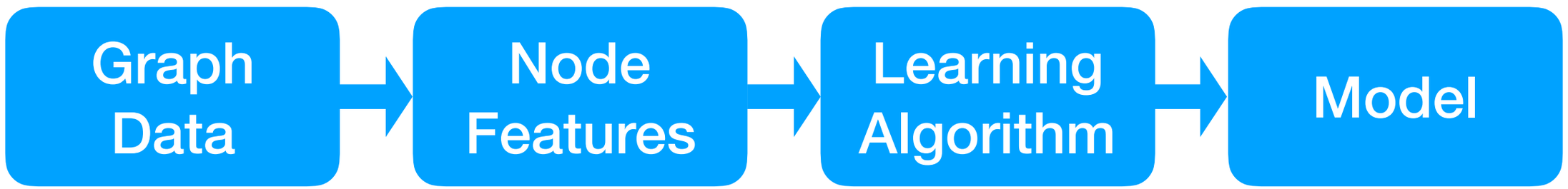
特征工程 Feature Engineering

Downstream Tasks下游应用

- Traditional social network analysis (SNA) measures structural properties of networks 传统社交网络分析考量图的结构属性
  - **Node**: degree, PageRank score, betweenness, closeness, eigenvector, …
    - 节点：度，pageRank分数，中介中心性，接近中心性，特诊向量等
  - **Pairs**: #common neighbors, ... 节点对：共同邻居节点等
  - **Groups**: cluster assignments, cliques, cores, clans, ... 团结构：集群分配，团，核等等。

图学习流程 Graph Learning Lifecycle

图数据 节点特征 学习算法 最终模型



特征工程
Feature Engineering

Downstream Tasks下游应用

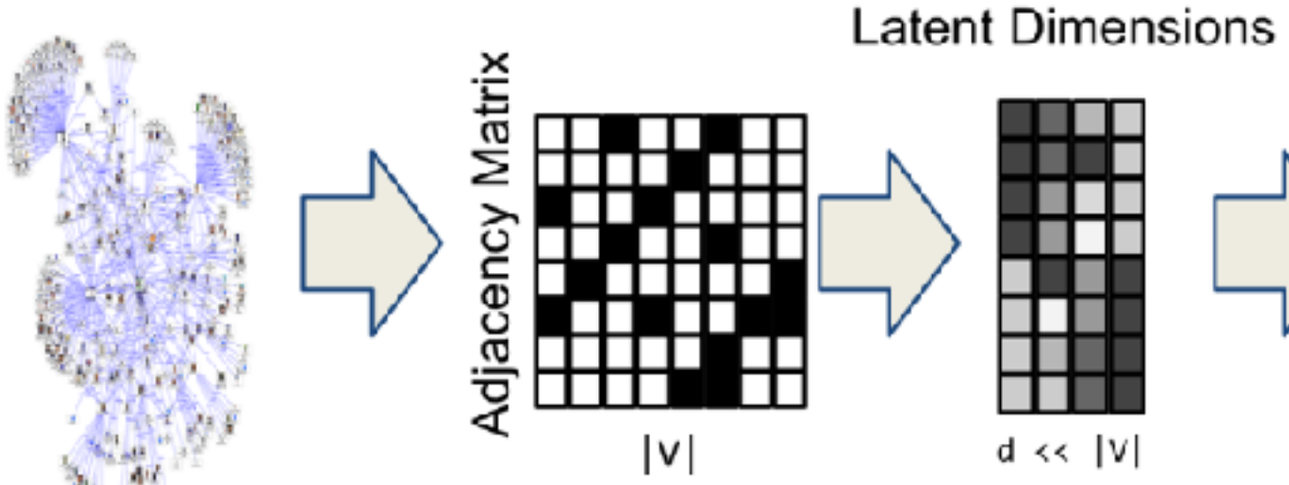
Feature Learning (implicit, distributed) 特征学习（隐形的，离散的）

- **DO NOT** need to compute the graph statistics as input node features：不需要用图节点特征计算图的统计数据
- Learning node features via **GRAPH EMBEDDING**： 通过图嵌入方法学习节点属性

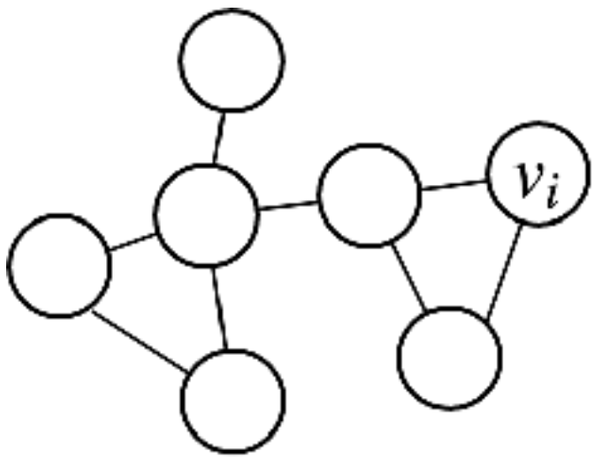
特征学习 Feature Learning

将图结构和节点属性映射到一个低维空间

- Map each node into **a low-dimensional space** by encoding **graph-structured** and **node attributes**



- Anomaly Detection 异常检测
- Attribute Prediction 属性预测
- Clustering 聚集
- Link Prediction 关系预测
- ...



$$f : v_i \rightarrow \mathbb{R}^d$$

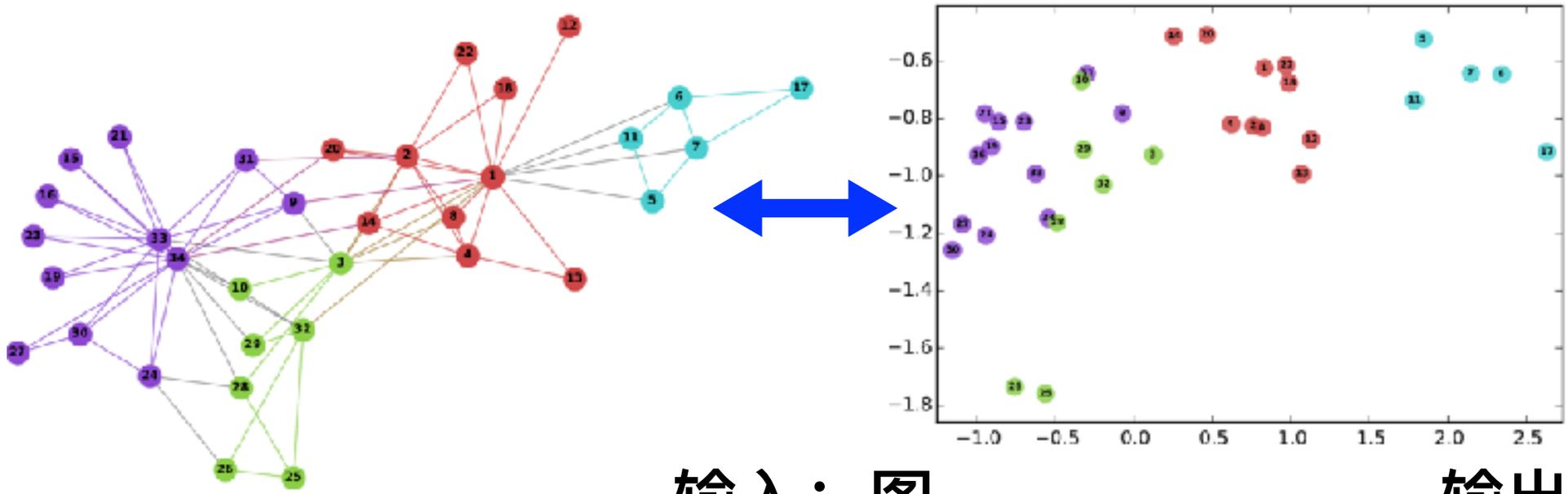
$$\mathbf{e}_i = [x_1^i, x_2^i, \ldots, x_d^i]$$

Node Vector / 点向量
Node Embedding / 点嵌入
点表示 Node Representation

举例

A Visual Example

- Zachary's Karate Club network: 空手道俱乐部网络



输入：图　　输出：图表示

(a) Input: Karate Graph　　(b) Output: Representation

[Perozzi et al. KDD 2014]

Intuition: Embedding of nodes to lower-dimension, so that "**similar**" nodes in the graph have embeddings that are **close** together

直觉：将节点嵌入到低维空间，所以语义相似的节点应在空间里距离相近

图嵌入

Graph Embedding

浅度嵌入方法

深度嵌入方法

Shallow Embedding Methods

Matrix Factorization Based Methods

基于矩阵分解

- [Belkin et al. NIPS 2002]
- [Ahmed et al. WWW 2013]
- [Cao et al. KDD 2015]
- [Ou et al. KDD 2016]

Random Walk Based Methods

基于随机漫步

- [Perozzi et al. KDD 2014]
- [Grover et al. KDD 2016]

Deep Embedding Methods

Recurrent Graph Neural Network

- [Li et al. ICLR 2015]
- [Dai et al. ICML 2016]

递归图神经网络

Spatial Temporal Graph Neural Network

- [Li et al. ICLR 2018]
- [Yan et al. AAAI'2018]
- [Wu et al. IJCAI'2019]
- [Guo et al. AAAI 2019]

时空图神经网络

Graph Autoencoder

图自动编码

- [Cao et al. AAAI 2016]
- [Wang et al. KDD 2016]

Graph Convolutional Network

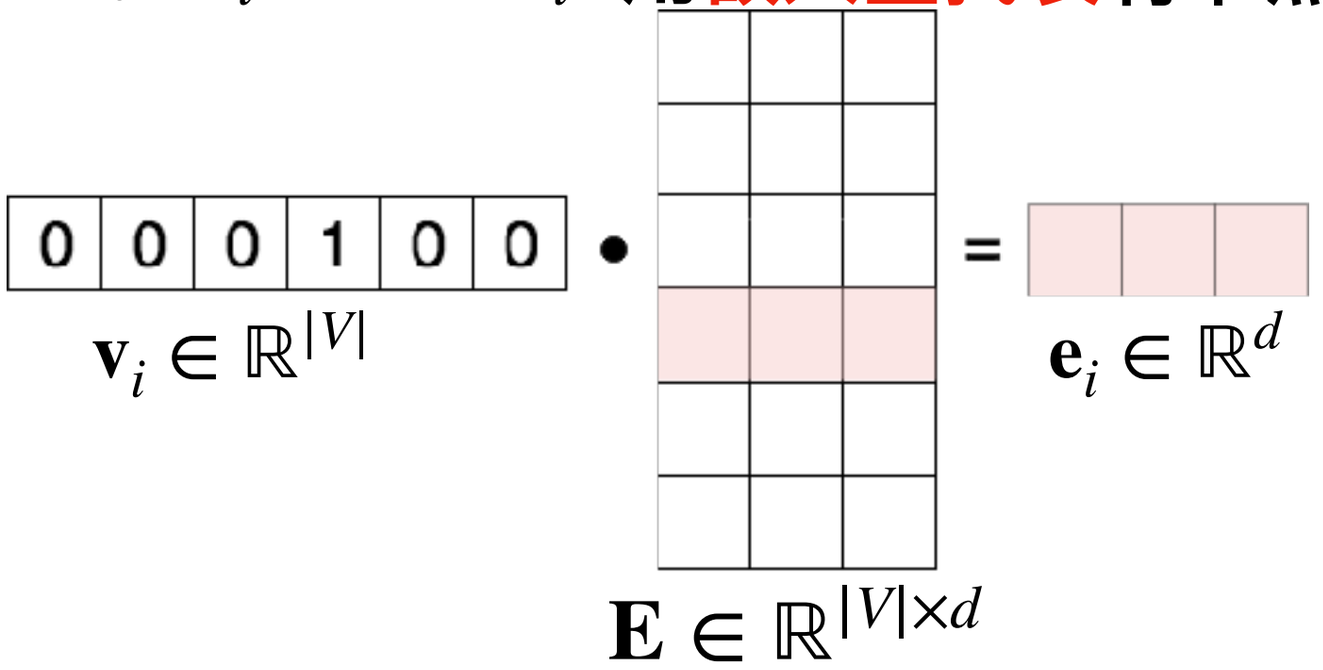
图卷积网络

- [Bruna et al., ICLR 2014]
- [Duvenaud et al., NIPS 2015]
- [Kipf & Welling, ICLR 2017]
- Our work [Zhang et al. UAI 2018]
- Our work [Zhang et al. IJCAI 2019]

Shallow Embedding Methods

浅度嵌入方法

Maps nodes to vector embeddings with an "**embedding lookup**" table: $f(\mathbf{v}_i, \mathbf{E}) = \mathbf{e}_i$ 用**嵌入查找表**将节点映射到矢量嵌入



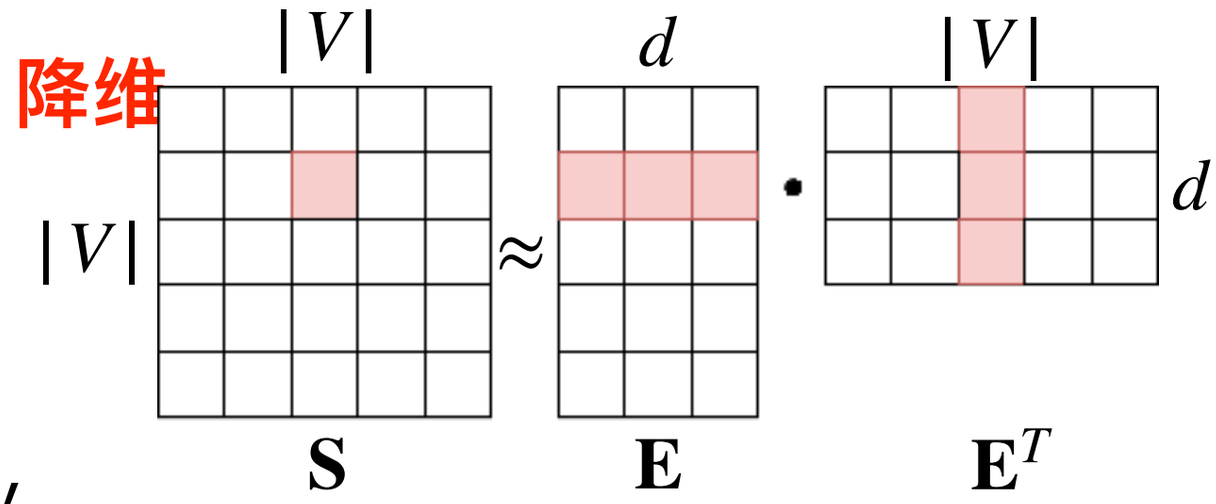
1. **Matrix Factorization-based Methods 基于矩阵分解方法**

2. **Random Walk-based Methods 基于随机漫步方法**

Matrix Factorization-Based

基于矩阵分解方法

- **Dimensionality Reduction 降维**



编码器

- **Encoder**: $f(\mathbf{v}_i) = \mathbf{v}_i\mathbf{E}$

解码器

- **Decoder**: pairwise similarity

损失

- **Loss**: $L \approx ||\mathbf{E}\mathbf{E}^T - \mathbf{S}||^2$

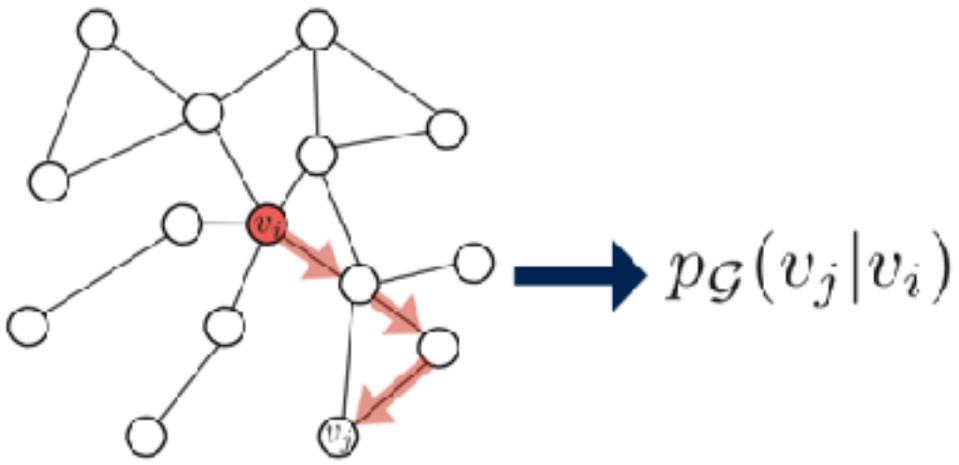
- $\mathbf{S}$ is a matrix of user-defined pairwise similarity measurement
  S是用户定义的衡量相似度的矩阵
- Laplacian Eigenmaps [Belkin et al. NIPS 2002]; Graph Factorization [Ahmed et al. WWW 2013]; GraRep [Cao et al. KDD 2015]; HOPE [Ou et al. KDD 2016]

Random Walk-Based

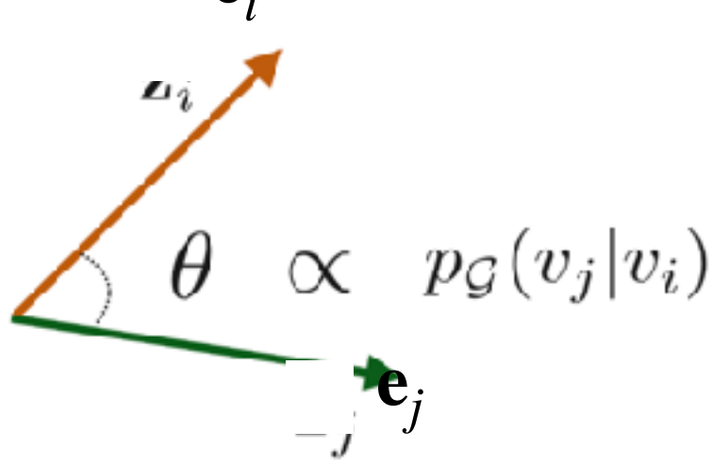
基于随机漫步方法

- Nodes **co-occur** on **short** random walks over the graph
  节点出现在短随机漫步的路径上

—> Nodes would have **similar** embeddings 节点有相似的嵌入



1. Run random walks to obtain co-occurrence statistics.



2. Optimize embeddings based on co-occurrence statistics.

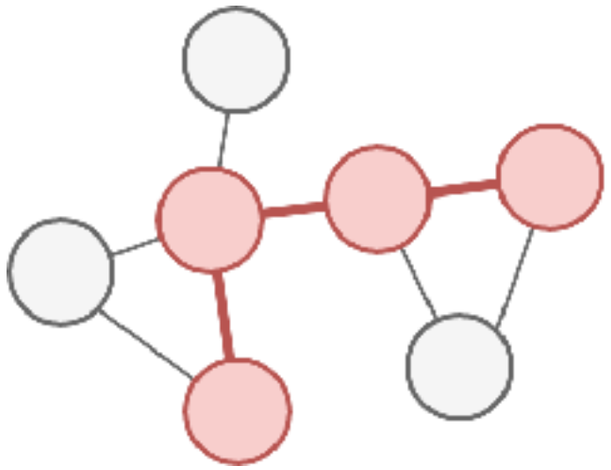
[Hamilton et al. IEEE 2018]

Random Walk-Based (Cont.)

基于随机漫步方法

- Deepwalk [Perozzi et al. KDD 2014] & Node2Vec [Grover et al. KDD 2016]

$$\mathbf{Dec}(\mathbf{e}_i, \mathbf{e}_j) \triangleq \frac{\mathbf{exp}(\mathbf{e}_i^T \mathbf{e}_j)}{\sum_{v_k \in V} \mathbf{exp}(\mathbf{e}_i^T \mathbf{e}_k)}$$



- Cross-entropy Loss: 交叉熵损失函数

$$L = \sum_{(v_i, v_j) \in D} -\mathbf{log}(\mathbf{Dec}(\mathbf{e}_i, \mathbf{e}_j))$$

Drawbacks of Shallow Embedding Methods

1. **No parameters sharing 无参数共享**
   - Simply an embedding lookup based on arbitrary node id 只是对任意点嵌入的查找

2. **Fail to leverage node attributes 无法充分利用节点属性**
   - E.g., user profiles on a social network 社交网络中用户的资料

3. **Cannot generate embeddings for previously unseen nodes 无法为新用户生成嵌入**
   - The Cold Start problem 冷启动问题

Deep Embedding Methods
基于深度嵌入方法

- Use more **complex** encoders 用更复杂的编码器
  - Often based on **deep neural networks 深度神经网络**
  - Depend more generally on the **structure** and **attributes** of the graph 更多依赖图的结构和属性

1. **Recurrent Graph Neural Network 递归图神经网络**
2. **Spatial Temporal Graph Neural Network 时空图神经网络**
3. **Graph Autoencoder 图自动编码**
4. **Graph Convolutional Network (GCN) 图卷积网络**

Recurrent Graph Neural Network

递归图神经网络

- Apply same set of parameters recurrently over nodes in a graph to extract high-level node representation. Node's hidden state is recurrently updated by
$h_v^{(t)} = \sum_{u \in N(v)} f(X_v, X_{(v,u)}^e, h_u^{(t-1)})$, where $f(\,.\,)$ is a parametric function

递归利用相同的参数在图上来训练获得点的向量表示

Same Graph Recurrent Layer (GREC) in updating node representations

时空图神经网络

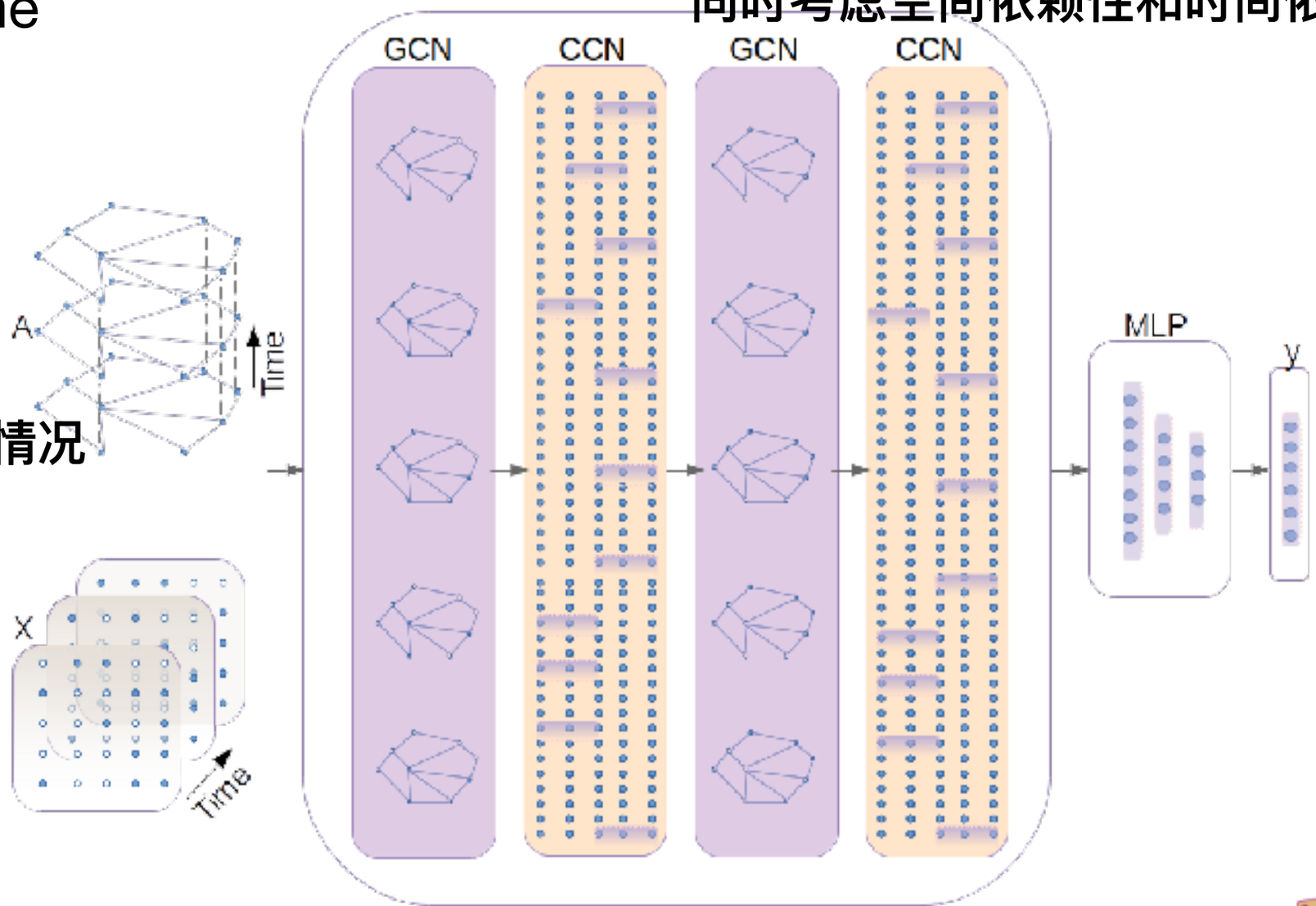
Spatial Temporal Graph Neural Network

- It considers spatial dependency and temporal dependency at the same time

同时考虑空间依赖性和时间依赖性

Eg. Spatial-temporal graph for traffic forecasting
[Yu et al. IJCAI,2018]

时空图以预测交通情况



时空图神经网络 Spatial Temporal Graph Neural Network

1. **STGNNs captures spatial and temporal dependencies of the graph simultaneously. STGNN 同时考虑了时间空间的依赖关系**

2. **The Task of STGNNs can be**
   - Forecasting future node values or labels 预测节点值或者标签等
   - Predicting spatial temporal graph labels 预测图的标签信息等等

3. **STGNN follow two directions**
   - RNN-based methods 基于循环神经网络模型的方法
   - CNN-based methods 基于卷积神经网络模型的方法

时空图神经网络 Spatial Temporal Graph Neural Network

- RNN-based approaches capture spatial-temporal dependencies by filtering inputs and hidden state passed to recurrent units using graph convolutions. For a simple RNN take $H^{(t)} = \sigma(WX^{(t)}) + UH^{(t-1)} + b)$,

After inserting graph convolution, above eq. becomes

$$H^{(t)} = \sigma(G_{conv}(X^{(t)}, A; W) + G_{conv}(H^{(t-1)}, A; U) + b)$$

通过过滤输入和传递给递归单元的隐藏层信息，
基于循环神经网络的方法综合考虑了时间空间依赖性

where H, W, U, X^t, b are hidden feature vector of node, weight vector of time step, weight vector for hidden layer, node feature vector at time t and dimension of H , respectively.

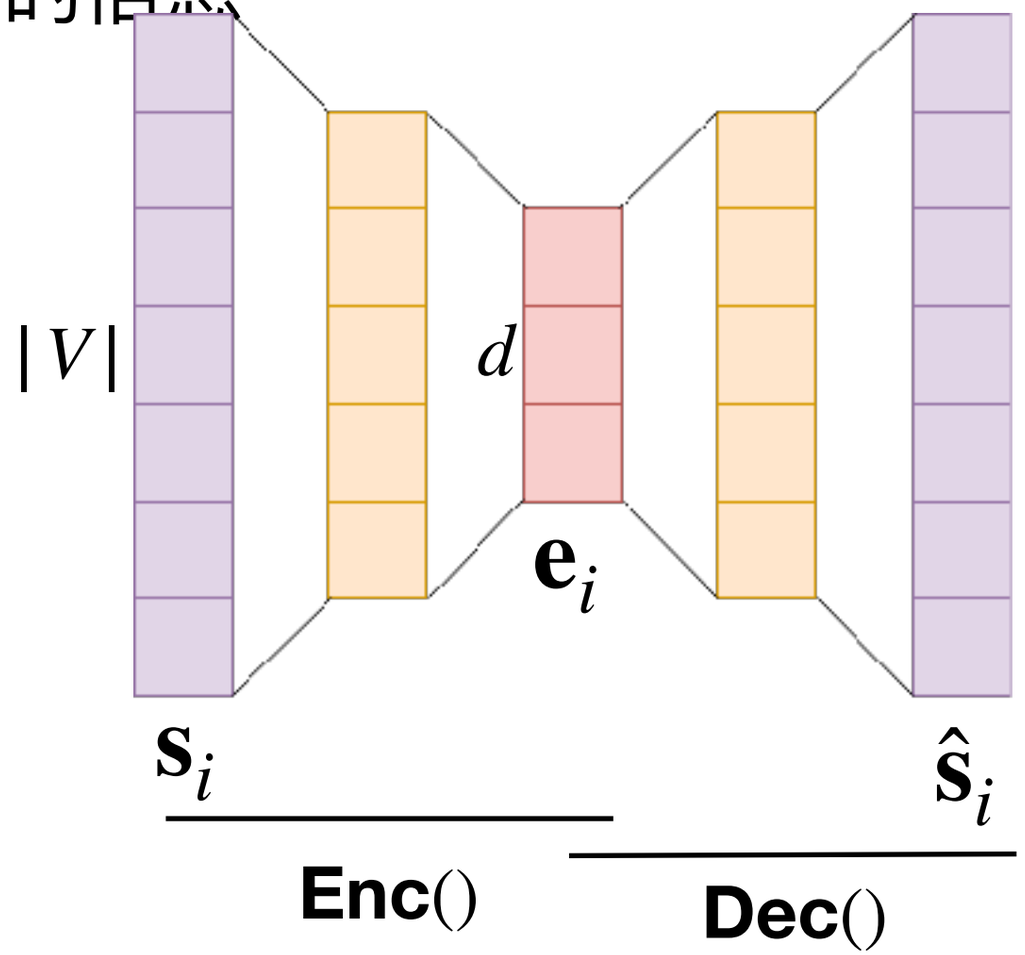
图自动编码 Graph Autoencoder

- To compress information about a node's local neighborhood 压缩点附近邻居的信息

$$\mathbf{Dec}(\mathbf{Enc}(\mathbf{s}_i)) = \mathbf{Dec}(\mathbf{e}_i) \approx \mathbf{s}_i,$$

$$L = \sum_{v_i \in V} ||\mathbf{Dec}(\mathbf{e}_i) - \mathbf{s}_i||^2$$

- $\mathbf{s}_i \in \mathbb{R}^{|V|}$
- $\mathbf{e}_i \in \mathbb{R}^d, d \ll |V|$



图自动编码 Graph Autoencoder

深度神经网络图表示

- Deep Neural Graph Representations [Cao et al. AAAI 2016]:
  - $\mathbf{s}_i \triangleq$ the pointwise mutual information of two nodes co-occurring on random walks

结构化深度网络嵌入

- Structural Deep Network Embeddings [Wang et al. KDD 2016]:
  - $\mathbf{s}_i \triangleq \mathbf{A}_i$ the adjacency vector of v_i

图自动编码 Graph Autoencoder

- **Advantages**:
  - Incorporate **structural** information as a form of **regularization**

    将结构信息纳入正则化
- **Disadvantages**:
  - Difficult to deal with **large scale** graphs 很难处理大规模图
    - The **input dimension** to the autoencoder is **fixed** at |V| 固定的输入维度
  - Cannot cope with **unseen nodes** 无法处理新节点
    - The **structure and size** of the autoencoder is **fixed** 结构和尺寸是固定的

Graph Convolutional Network (GCN)
图卷积网络

- Spatial Approach:基于空间信息
  - Aggregate information from **Local Neighborhood**

综合考虑局部邻居信息和参数共享 + Parameter Sharing

1. Define **neighborhood**
   - All neighbors  **所有邻居**
   - Fixed size uniform sampling  **固定尺寸统一采样**
   - Random walk sampling  **随机漫步采样**
2. Design graph **aggregator** $\mathbf{y}_i = r_\theta(\mathbf{x}_i, \{\mathbf{z}_i\})$ **设计图聚集器**

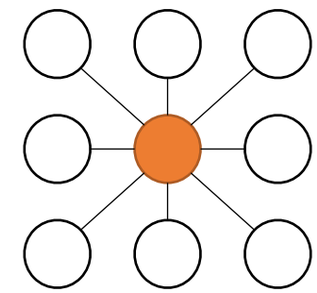
Graph Aggregator Properties
图聚集器性质

Patch --> Regular Grid 规则网络

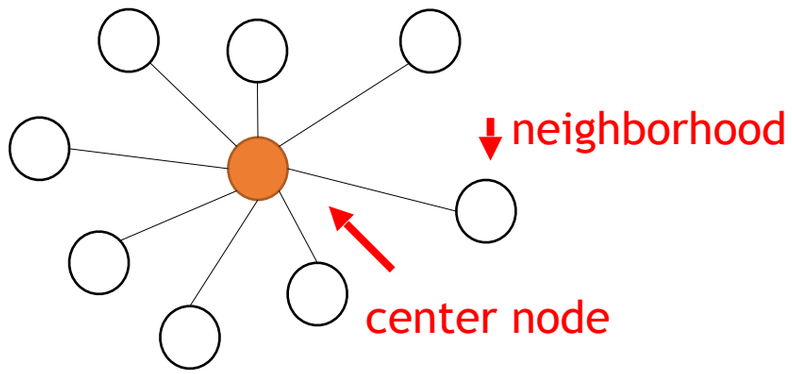
1. Permutation-sensitive
   排列方式：敏感
2. Fixed Size
   固定尺寸

Set --> Irregular Grid 不规则网络

1. **Permutation-invariant**
   排列方式：恒定的
2. **Dynamic Resize**
   动态尺寸

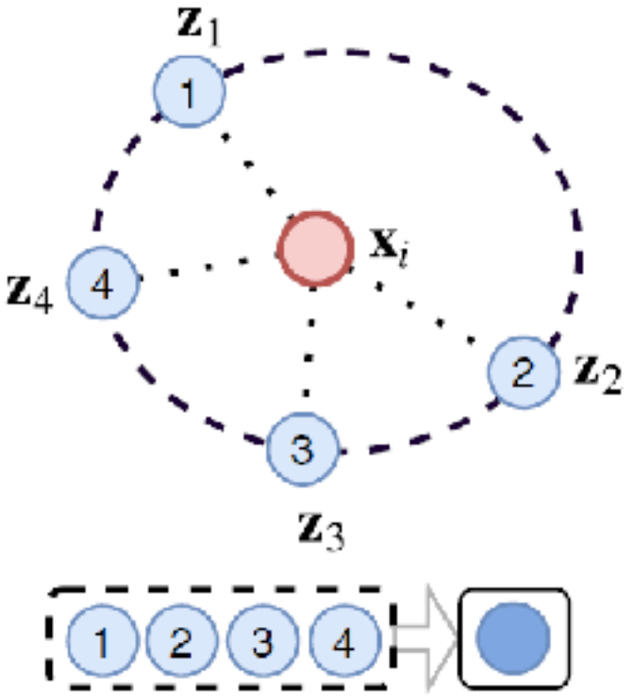


Convolution

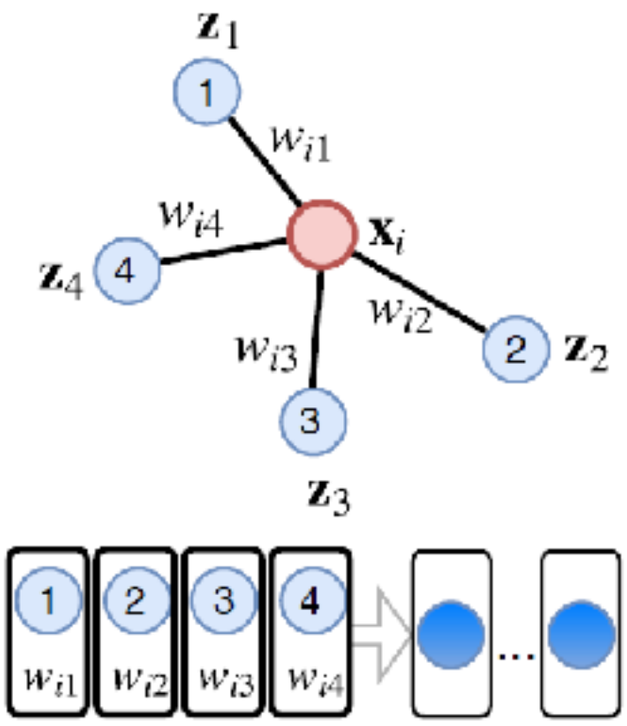


Graph Convolution

图聚集器 Graph Aggregators



Pooling-based



Pairwise-sum

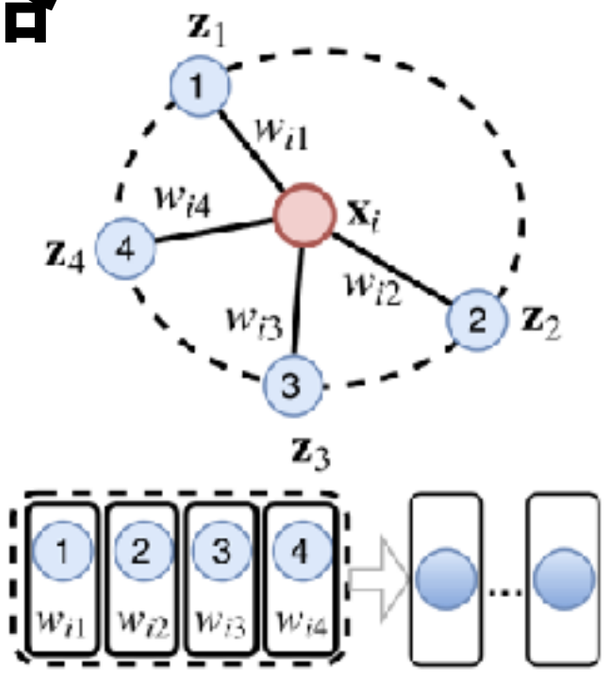
$$\mathbf{y}_i = \phi_o(\mathbf{x}_i \oplus \text{pool}_{j \in \mathcal{N}_i}(\phi_v(\mathbf{z}_j)))$$

$$\mathbf{y}_i = \phi_o(\mathbf{x}_i \oplus \Big\|_{k=1}^{K} \sum_{j \in \mathcal{N}_i} w_{i,j}^{(k)} \phi_v^{(k)}(\mathbf{z}_j)),$$

$$w_{i,j}^{(k)} = \phi_w^{(k)}(\mathbf{x}_i, \mathbf{z}_j).$$

Graph Aggregators (Cont.)

图聚集器



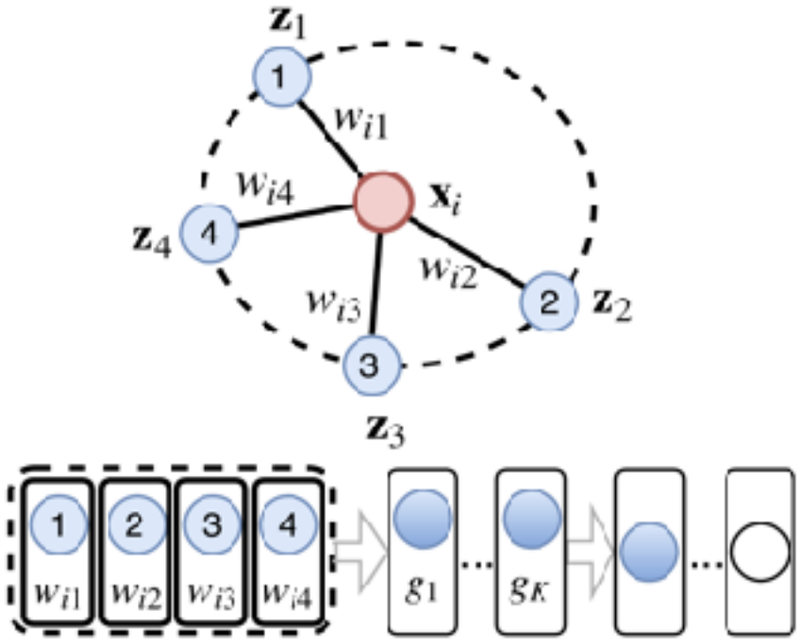
Attention-based Aggregator

基于注意力机制的聚集器

$$\mathbf{y}_i = \mathrm{FC}_{\theta_o}(\mathbf{x}_i \oplus \Big\|_{k=1}^{} \sum_{j \in \mathcal{N}_i} w_{i,j}^{(k)} \mathrm{FC}_{\theta_v^{(k)}}^{h}(\mathbf{z}_j)),$$

$$w_{i,j}^{(k)} = \frac{\exp(\phi_w^{(k)}(\mathbf{x}_i, \mathbf{z}_j))}{\sum_{l=1}^{|\mathcal{N}_i|} \exp(\phi_w^{(k)}(\mathbf{x}_i, \mathbf{z}_l))},$$

$$\phi_w^{(k)}(\mathbf{x}, \mathbf{z}) = \langle \mathrm{FC}_{\theta_{xa}^{(k)}}(\mathbf{x}), \mathrm{FC}_{\theta_{za}^{(k)}}(\mathbf{z}) \rangle.$$



Gated Attention Aggregator (GaAN)

基于门控注意力机制的聚集器

lgate

$$\mathbf{y}_i = \mathrm{FC}_{\theta_o}(\mathbf{x}_i \oplus \Big\|_{k=1}^{K} (g_i^{(k)} \sum_{j \in \mathcal{N}_i} w_{i,j}^{(k)} \mathrm{FC}_{\theta_v^{(k)}}^{h}(\mathbf{z}_j))),$$

$$\mathbf{g}_i = [g_i^{(1)}, ..., g_i^{(K)}] = \psi_g(\mathbf{x}_i, \mathbf{z}_{\mathcal{N}_i}),$$

[Zhang et al. UAI 2018]

Graph Neural Networks for Recommendation

图聚集器

- Formalize the user-item interaction data as a **bipartite graph**

  将用户和物品的交互关系描述成二部图

- **Rating prediction**

  评分预测

  - Predict some missing ratings given the existing rating pairs

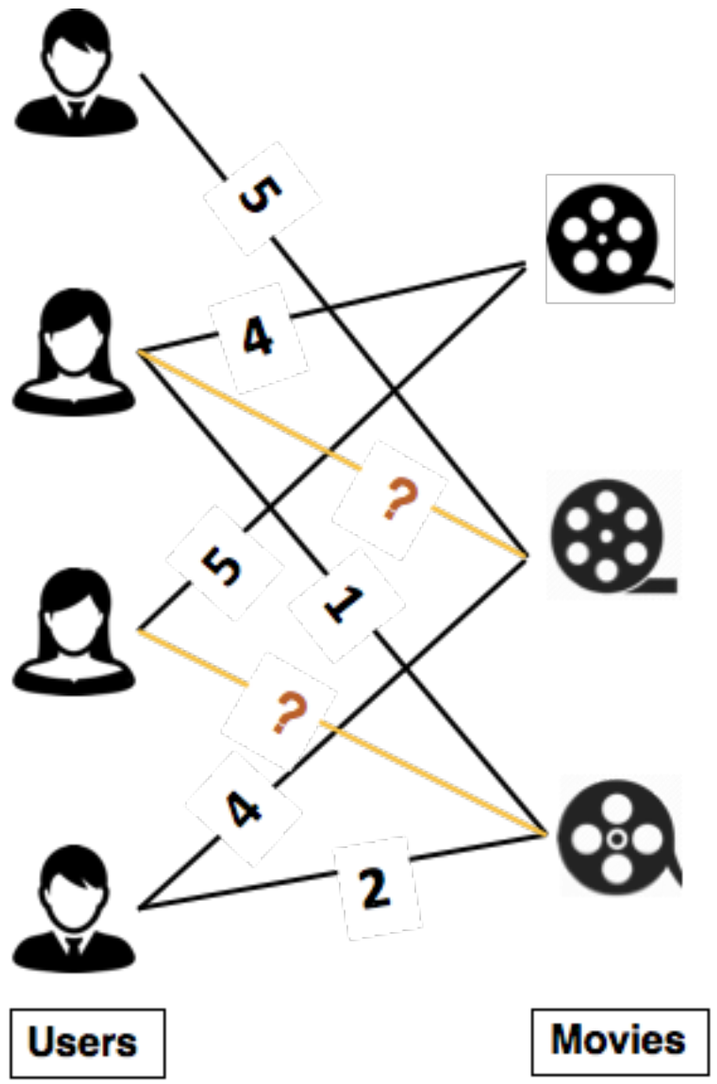
  基于已有的评分信息预测缺省的评分

- **Transductive** rating prediction

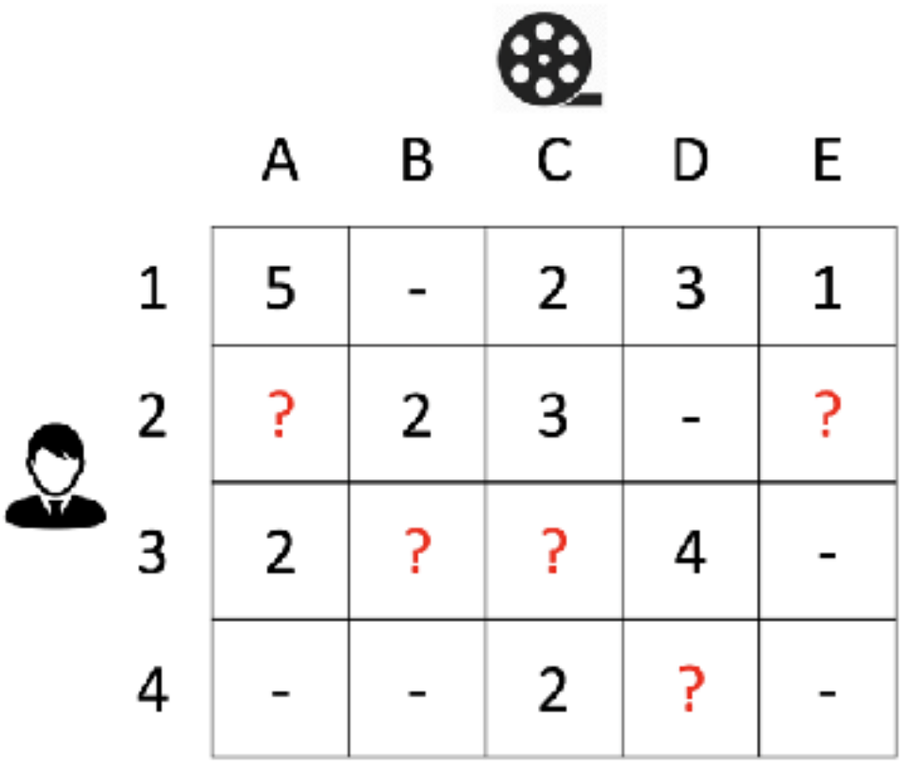
  转导评分预测

- **Cold start** scenario

  冷启动场景



Transductive Rating Prediction
转导评分预测



- Train on observed ratings
  在可观测评分上训练
- Predict the missing ratings , i.e., `?`
  预测缺省评分
- All the testing users and items are **observed** in training data

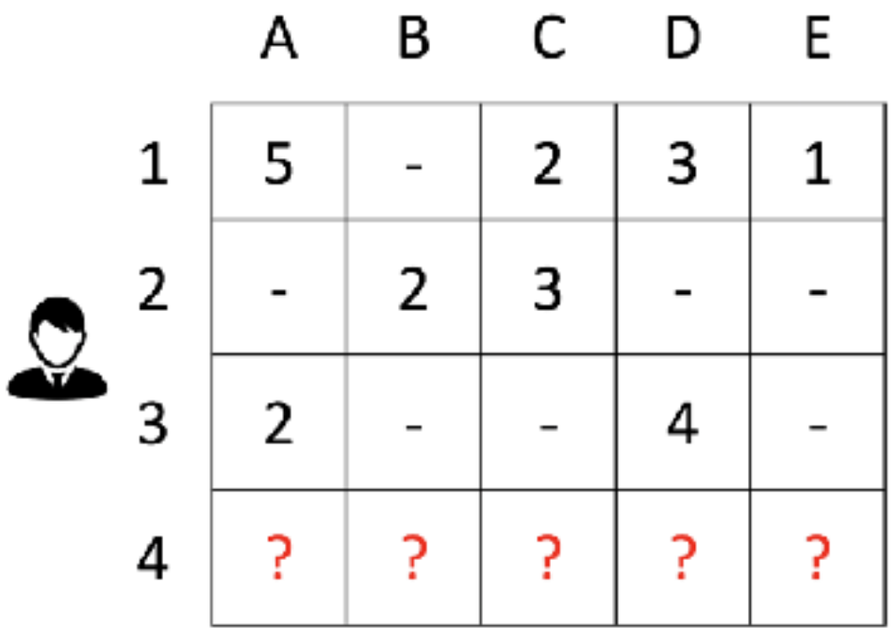
在训练数据集中所有测试用户和物品均可以观测

- **Matrix completion** problem
  - Matrix Factorization (MF) 矩阵分解

矩阵补全问题

Cold Start Scenario 1: New Users/Items

冷启动示例1 新用户/物品

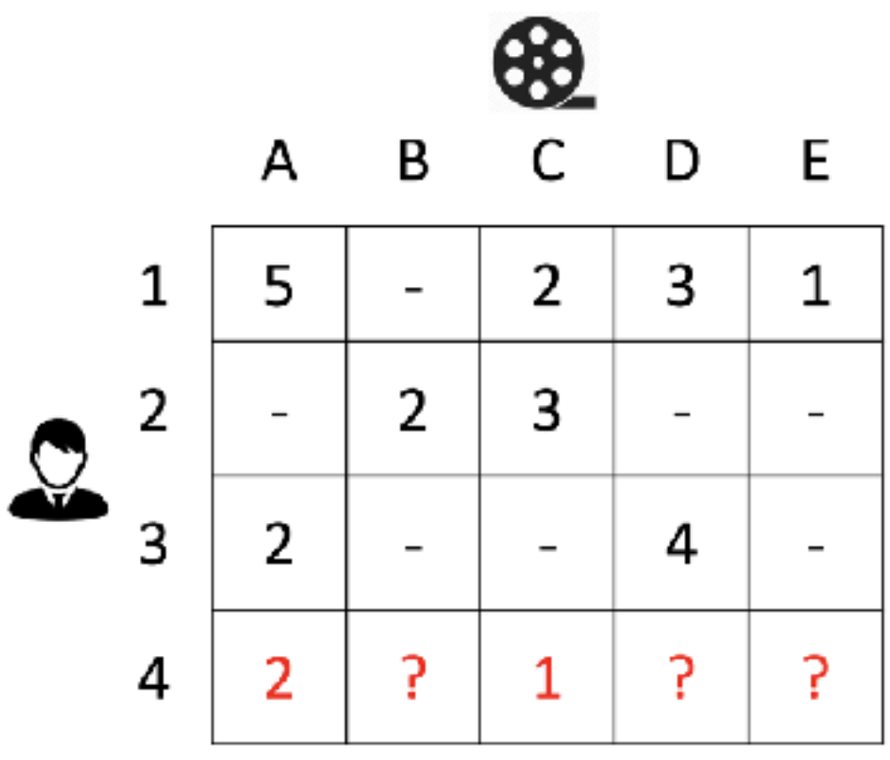


	A	B	C	D	E
1	5	-	2	3	1
2	-	2	3	-	-
3	2	-	-	4	-
4	?	?	?	?	?

- Train on observed ratings
  在可观测评分数据上训练
- Predict how **a new user** will rate the movie
  预测新用户如何给电影评分
- Content-based recommendation 基于内容的推荐
  基于协同深度学习的推荐
  - Collaborative Deep Learning for Recommender Systems
    [Wang et al. KDD 2015]
  - To encode the user features, e.g., profiles
    将用户属性进行编码

Cold Start Scenario 2: Ask-to-rate

冷启动示例2 邀请评分



	A	B	C	D	E
1	5	-	2	3	1
2	-	2	3	-	-
3	2	-	-	4	-
4	2	?	1	?	?

- Train on the observed ratings

在可观测评分上训练

- Test time:
  - Ask **a new user** to **rate several movies**

邀请新用户对电影评分

  - Predict how a user will rate other movies

预测用户如何评价其他电影

- **Inductive** rating prediction

归纳评分预测

Stacked and Reconstructed Graph Convolutional Networks

堆叠重构图卷积网络

[Zhang et al. IJCAI 2019]

- **STAR-GCN**:
  - Can solve the **cold start problem** (**inductive**) in recommender systems

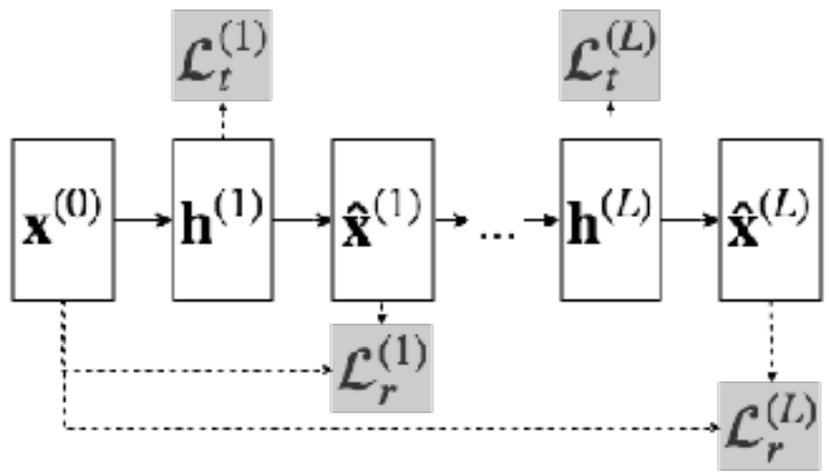
    将节点嵌入到低维向量 可处理大规模图
    - Embed nodes to low-dimensional vectors -> Scalable to large graphs
    - Multi-block structure: Mask and reconstruct node embeddings (BERT [Jacob et al. 2019] for Graph) **多区块结构：隐藏和重构节点嵌入**

- **SOTA results** on both **transductive** and **inductive** (ask-to-rate) rating prediction task
  - Achieves the **best state-of-the-art** results

    在转导和归纳评分预测问题上取得了目前最优结果

STAR-GCN

堆叠重构图卷积网络



多区块图编码解码

- A multi-block graph encoder-decoder

$$\mathbf{x}^{(0)} \rightarrow \mathbf{h}^{(1)} \rightarrow \hat{\mathbf{x}}^{(1)} \rightarrow \mathbf{h}^{(2)} \rightarrow \hat{\mathbf{x}}^{(2)} \rightarrow \ldots \rightarrow \hat{\mathbf{x}}^{(L)}$$

- Go beyond BERT
  - BERT has a single block $\mathbf{x}^{(0)} \rightarrow \mathbf{h}^{(1)} \rightarrow \hat{\mathbf{x}}^{(1)}$

通过编码语义图结构和特征输入生成节点表示

- Encoder ($\mathbf{x}^{(l-1)} \rightarrow \mathbf{h}^{(l)}$)
  - Generate node representations by encoding semantic graph structures + input features
- Decoder ($\mathbf{h}^{(l)} \rightarrow \hat{\mathbf{x}}^{(l)}$)
  - Recover masked input node embeddings 恢复隐藏的输入节点嵌入
  - Any variant of graph convolutional network (GCN) can be an encoder or decoder
- Loss 评分预测损失函数
  - A rating prediction loss $\mathcal{L}_t$
  - A node reconstruction loss $\mathcal{L}_r$

节点重构损失函数

GCN的任何变体可以是一个编码器或者解码器

Experiment: Dataset

Table 1: Statistics of the datasets. 'D_U' and 'D_V' are the input feature dimension of users and items, respectively.

	D_U	D_V	#U	#V	$\mathcal{R}$	#R
Flixster	3K	3K	2,341	2,956	0.5,1,..,5	26,173
Douban	3K	-	2,999	3,000	1,..,5	136,891
ML-100K	23	320	943	1,682	1,..,5	100K
ML-1M	23	320	6,040	3,706	1,..,5	1M
ML-10M	-	321	69,878	10,677	0.5,1,..,5	10M

Experiments: Transductive Rating Prediction

	Flixster	Douban	ML-100K	ML-1M	ML-10M
BiasMF [Koren *et al.*, 2009]	-	-	0.917	0.845	0.803
NNMF [Dziugaite and Roy, 2015]	-	-	0.907	0.843	-
I-AUTOREC [Sedhain *et al.*, 2015]	-	-	-	0.831	0.782
GRALS [Rao *et al.*, 2015]	1.245	0.833	0.945	-	-
CF-NADE [Zheng *et al.*, 2016]	-	-	-	**0.829**	0.771
Factorized EAE [Hartford *et al.*, 2018]	-	-	0.910	0.860	-
sRMGCNN [Monti *et al.*, 2017]	0.926	0.801	0.929	-	-
GC-MC [Berg *et al.*, 2017]	0.917	0.734	0.910	0.832	0.777
STAR-GCN	**0.879**±0.0030	**0.727**±0.0006	**0.895**±0.0009	0.832±0.0016	**0.770**±0.0001

RMSE: the lower the score, the better the performance

- **STAR-GCN** architecture achieves the **best state-of-the-art** results on **four out of five** datasets

Star-GCN在四个数据集上获得了最佳结果

Experiment: Inductive Rating Prediction (Ask-to-rate)

目的：预测在训练阶段用户或者物品缺失的评分信息

- **Goal**: to predict ratings where either users or items are not seen in the training phase
- Experiment Setting:
  - Keep 20% of user (or item) nodes as the testing nodes & remove them from the training graph

Datasets	Models	Items 20%			Users 20%		
		50%	30%	10%	50%	30%	10%
Douban	DropoutNet	-	-	-	0.797±0.002	0.797±0.003	0.797±0.001
	CDL	-	-	-	0.781±0.006	0.781±0.001	0.781±0.001
	STAR-GCN(- rec.)	0.734±0.001	0.746±0.001	0.777±0.002	0.731±0.000	0.738±0.000	0.753±0.001
	STAR-GCN(- rec., + fea.)	-	-	-	0.731±0.002	0.737±0.000	0.753±0.001
	STAR-GCN	**0.725**±0.001	**0.734**±0.001	**0.764**±0.000	**0.725**±0.001	**0.731**±0.001	0.747±0.001
	STAR-GCN(+ fea.)	-	-	-	**0.725**±0.002	**0.731**±0.000	**0.746**±0.000
ML-100K	DropoutNet	1.223±0.065	1.167±0.031	1.144±0.024	1.015±0.002	1.022±0.006	1.023±0.003
	CDL	1.083±0.009	1.082±0.007	1.082±0.007	1.011±0.005	1.013±0.006	1.015±0.004
	STAR-GCN(- rec.)	0.932±0.001	0.943±0.001	0.976±0.003	0.919±0.002	0.933±0.001	0.949±0.001
	STAR-GCN(- rec., + fea.)	0.928±0.002	0.941±0.002	0.977±0.004	0.916±0.005	0.931±0.004	0.951±0.005
	STAR-GCN	0.919±0.001	**0.926**±0.000	**0.954**±0.001	**0.907**±0.004	**0.917**±0.005	0.937±0.005
	STAR-GCN(+ fea.)	**0.918**±0.002	**0.926**±0.002	0.956±0.000	**0.907**±0.002	**0.917**±0.001	**0.936**±0.004
ML-1M	DropoutNet	1.169±0.120	1.134±0.034	1.256±0.128	1.002±0.001	1.005±0.005	1.003±0.001
	CDL	1.068±0.009	1.069±0.009	1.068±0.009	0.974±0.000	0.975±0.000	0.974±0.000
	STAR-GCN(- rec.)	0.862±0.001	0.872±0.004	0.903±0.004	0.859±0.002	0.868±0.001	0.891±0.001
	STAR-GCN(- rec., + fea.)	0.861±0.002	0.867±0.002	0.910±0.006	0.859±0.001	0.869±0.001	0.893±0.001
	STAR-GCN	**0.844**±0.000	**0.850**±0.000	**0.876**±0.004	**0.848**±0.001	**0.858**±0.001	**0.882**±0.000
	STAR-GCN(+ fea.)	**0.844**±0.001	0.851±0.001	**0.876**±0.002	0.849±0.001	**0.858**±0.000	0.883±0.001

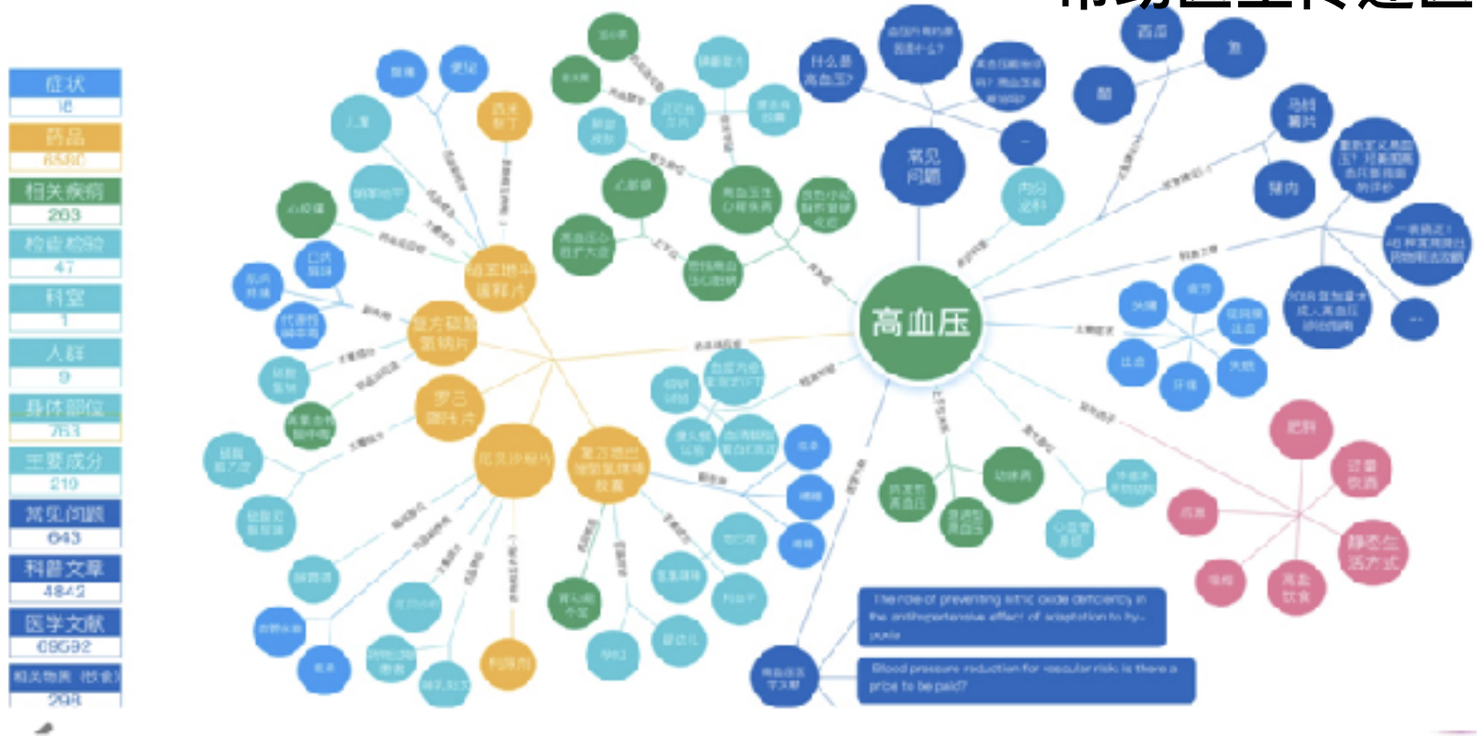
RMSE: the lower the score, the better the performance

- STAR-GCN produces **significantly better** results than baselines

医学图嵌入
Medical Graph Embedding

Medical Knowledge Graph 医学知识图谱

- including medicine, disease, symptoms, electronic medical records, etc... 药物，疾病，症状，电子病历等等
  节点之间的关系代表医学知识
- Internal edges representing medical knowledges.
- Help doctors to deliver the knowledge and decision support
  帮助医生传递医学知识和决策支持



医学图嵌入
Medical Graph Embedding

- **Research Question : How to do graph embedding for medical knowledge graphs?**

- **Observation: A patient with disease "pneumonia" has symptom "fever".**
  观察：有肺炎的病人有发烧的症状
- Conventional model: "pneumonia" leads to "fever".
  传统方法：肺炎导致发烧
- Our thinking: "fever" is common but not always presented in disease "pneumonia"!
  我们认为：发烧并不一定由肺炎导致
- *P(symptom = fever | disease = pneumonia)*

- **We are doing: probabilistic graph embedding for medical knowledge graphs** **我们用概率图嵌入来表达医学知识图谱**

Drug-Drug Interaction
药物反应预测Prediction (DDI)

- Motivation: Clinical studies cannot sufficiently and accurately identify DDI's.
  动机：临床试验无法充分检测药物反应
- How about using AI to help predict DDI's?
  如何用AI帮助预测药物反应?
- **We are doing: using medical knowledge graphs to predict DDI's through link prediction.**

我们将药物反应问题转化为医学知识图谱的关系预测问题

Sampling Large Graphs

大规模图问题

Large Graph issues

- **Ubiquity of large graph composed of millions of nodes and edges　百万节点规模图很常见**
- **In order to study it we require to store and compute the whole graph 为了学习大规模图，我们需要存储和计算**
- **It raises space and computation issues even to compute basic properties of the graph**
  即便是计算图的基础属性空间和时间复杂度都被提高了

Research Question : Given a large real graph, how can we derive a representative sample preserving properties of original graph?

研究问题：如何在大规模图上实现具有代表性的采样?

Sampling Large Graphs

大规模图问题

Challenges

现有算法对大规模图效率低

- **Current algorithms are slow for large graph**
- **No optimal solution for graph sampling** 未有图采样的最优算法
- **Existing sampling approaches can only target one criteria / property to match with original graph**
  现有采样算法只能针对一个特征进行采样
- **No baseline machine learning solution to graph sampling**
  未出现用于图采样的机器学习算法

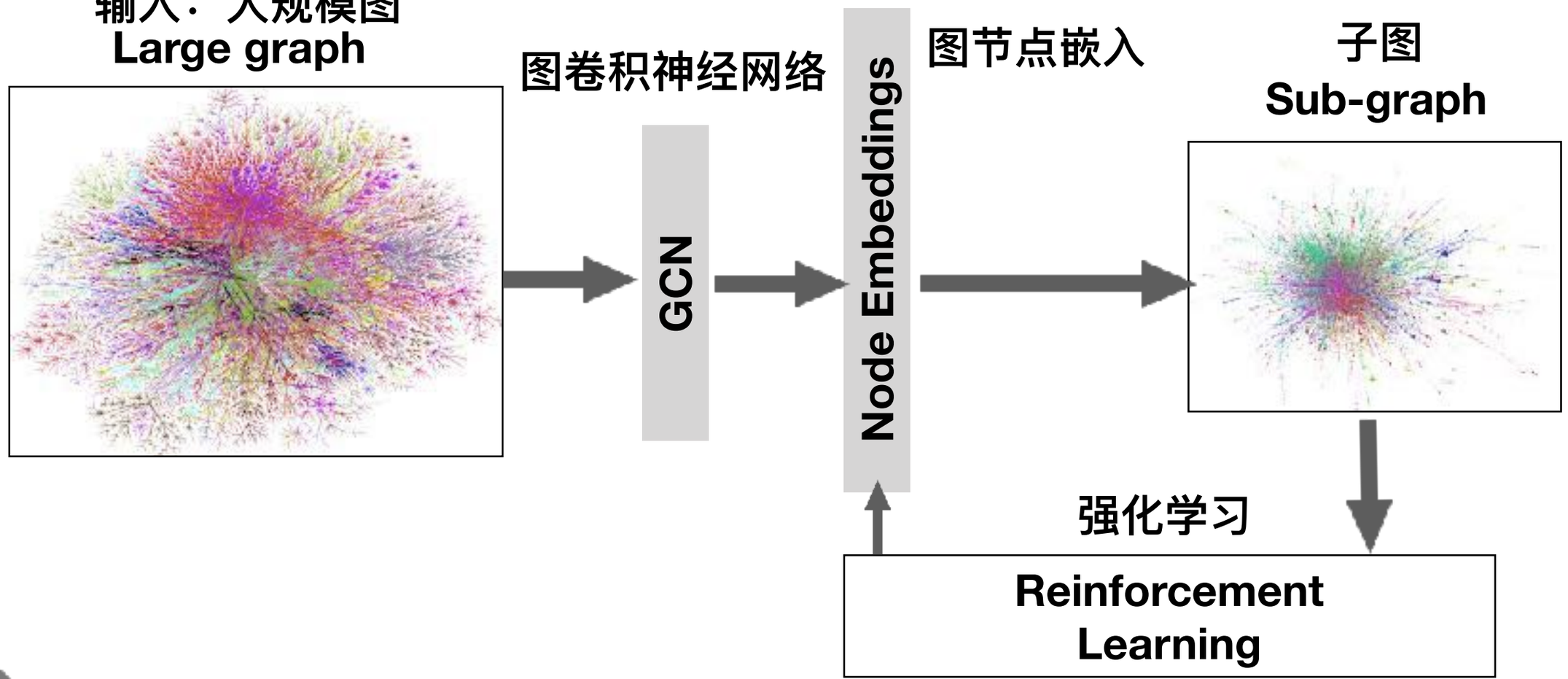
Sampling Large Graphs

大规模图问题

Proposed Solution :

利用深度强化学习进行图采样

Graph sampling using deep reinforcement learning



Summary & Future Directions

- Graph embedding (Graph Convolution Network) is a **powerful approach** to obtain **implicit** and **distributed** representation of a graph with **node (attributes)** and **link (relations)** information!

  GCN图嵌入是一个可获取节点、关系隐性和分布式表示的强大的方法

- **STAR-GCN** achieves SOTA results in **transductive and inductive rating prediction** [Zhang et al. UAI 2018]

  STAR-GCN 在转导和归纳评分预测任务上获得了最佳结果

- **Gated Attention Networks** achieve SOTA results in **inductive node classification** [Zhang et al. IJCAI 2019]

  门控注意力网络在归纳节点分类问题是获得了最佳结果

- Future Directions

  时间和异质图神经网络

  - **Temporal** and **heterogenous** graph neural networks
  - **Quantum random walk** algorithm for neighborhood sampling

  局部采样的量子随机漫步算法

Source Code

- Code: https://github.com/jennyzhang0215/STAR-GCN

- [**Zhang et al. UAI 2018**] Zhang, Jiani, Xingjian Shi, Junyuan Xie, Hao Ma, Irwin King, and Dit-Yan Yeung. "*GAAN: Gated attention networks for learning on large and spatiotemporal graphs.*" UAI 2018.

- [**Zhang et al. IJCAI 2019**] Zhang, Jiani, Xingjian Shi, Shenglin Zhao, and Irwin King. "*STAR-GCN: Stacked and Reconstructed Graph Convolutional Networks for Recommender Systems*." IJCAI 2019.

THE WEB CONFERENCE | April 20-24 TAIPEI 2020

https://www2020.thewebconf.org

April 20-24, 2020

ABOUT AUTHORS ATTENDEES PROGRAMS SPONSORS

The Chinese University of Hong Kong

Thanks!