

日期： /

Fast Neural Subgraph Matching and Clustering

[Node-included subgraph (节点诱导子图)]

给定图 $G = (V, E)$, 节点诱导子图 $G'(V', E')$ 通过以下方式定义

$$\cdot V' \subseteq V$$

$$\cdot E' = \{(u, v) \in E \mid u, v \in V'\}$$

G' 是由 V' 诱导的 G 的子图 (化学, 分子表示.)

[Edge-induced subgraph (边诱导子图/非诱导子图)]

边诱导子图与结点诱导子图类似：

$$\cdot E' \subseteq E$$

$$\cdot V' = \{v \in V \mid (u, v) \in E' \text{ for some } u\}$$

$G'(V', E')$ 是 $G(V, E)$ 的, E' 诱导的子图 (知识图谱)

如何确定图与子图的包含关系?

[Graph isomorphism problem (图同构问题)]

当图 $G_1(V_1, E_1)$ 和 $G_2(V_2, E_2)$ 之间可以建立单射 $f: V_1 \rightarrow V_2$ 使得

$\forall (u, v) \in E_1$ 有 $(f(u), f(v)) \in E_2$ 时, 称 G_1, G_2 同构

目前没有用于图同构问题的多项式时间算法。

当 G_2 的部分子图与 G_1 同构时, 称 G_2 是 G_1 的子图同构。

日期： /

[Network motifs (网络模体)]

在图中反复出现的具有显著连接性的模式。

[子图频率 (subgraph frequency)]

1) 图级别 (Graph-level)

G_Q 在 G_T 中的频率，定义为 G_T 中可以诱导为与 G_Q 同构的子图的独特节点集合数。

2) 节点级别 (Node-level)

与 Graph-level 的定义不同，Node-level 的定义要求在 G_Q 中选择一个锚点，而 G_Q 在 G_T 中的频率定义为 G_T 中 节点 u 的个数。
注意是节点个数
而非同构子图的
个数

Q：如何处理含多个图的数据集？

A：将各图视为一个非连通的，含多个连通分量的大图的子图。

[Motif Significance]

如何生成随机图

我们认为在真实图中出现频率显著高于随机图的 Motif 有高 significance。

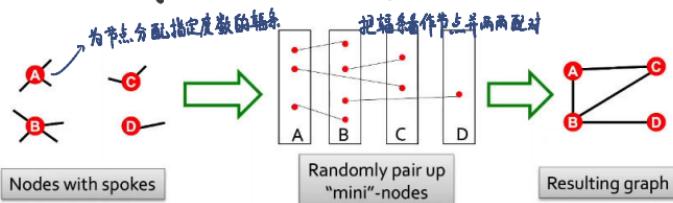
日期: /

[Erdős - Rényi (ER) random graphs]:

$G_{n,p}$ 定义了一个包含 n 个节点的图，各节点连接的概率为 p

[Configuration Model]

目标是生成具有给定度序列 $k_1, k_2 \dots k_N$ 的随机图



[Z-score]

给定子图或模体 i 的 Z-score 捕获了其在真实图和随机图中分布差异的统计显著性。

$$Z_i = (N_i^{\text{real}} - \bar{N}_i^{\text{rand}}) / \text{std}(N_i^{\text{rand}})$$

有值图出现的平均频数

当 $Z > 2.2$ 时认为频率差异显著。

定义网络显著性概况 (Network significance profile / SP)

$$SP_i = Z_i / \sqrt{\sum Z_i^2} \rightarrow \text{含相同节点数的子图/模体的 Z-score}$$