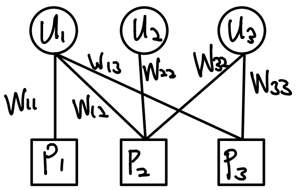
undirected bipartite



U, P — vertex sets

E — edge sets (bipartite — only between vertices of different types)

$|U|$ ($|P|$) — number of elements in U (P)

w_{ij} — non-negative, relationship strength between u_i and p_j (= 0 if not connected)

W — $|U| \times |P|$ matrix, $= [w_{ij}]$

d_i — u_i's weighted degree (sum of connected edges' weights)

D_u (D_p) — diagonal matrix, weighted degrees of vertices in U (P). $(D_u)_{ii} = d_i$ ($(D_p)_{jj} = d_j$)

query vector — prior belief of vertices with respect to the ranking criterion
(since ranking simply based on the graph structure is insufficient)

Input — $G = (U \cup P, E)$, weight matrix W, query vector $\mathbf{u}^0, \mathbf{p}^0$

Output — $f: P \cup U \rightarrow \mathbb{R}$ (maps each vertex to a real number)

$f(u_i)$ ($f(p_j)$) — ranking score, also denotes u_i (p_j)

final ranking score ⇒ $\mathbf{u} = [u_i]$, $\mathbf{p} = [p_j]$ (vectors)

Design

- follow smoothness convention: a vertice should be ranked high if connected to higher-ranked vertices

⇒ defines a mutually-reinforcing relationship

$$p_j = \sum_{i=1}^{|U|} w_{ij} u_i \qquad u_i = \sum_{j=1}^{|P|} w_{ij} p_j$$

additive update rule

⇒ normalization is necessary (ensure stability & convergence)

☆ - Birank adopt the symmetric normalization scheme

⇒ smooth an edge weight by the degree of its two connected vertices simultaneously

$$p_j = \sum_{i=1}^{|U|} \frac{w_{ij}}{\sqrt{d_i}\sqrt{d_j}} u_i \qquad u_i = \sum_{j=1}^{|P|} \frac{w_{ij}}{\sqrt{d_i}\sqrt{d_j}} p_j$$

⇒ allowing edges connected to high-degree vertex to be suppressed, lessening the contribution of high-degree vertices

– factor the query vector directly into the ranking process (post-processing doesn't work in some context)

$$p_j = \alpha \sum_{i=1}^{|U|} \frac{w_{ij}}{\sqrt{d_i}\sqrt{d_j}} u_i + (1-\alpha) p_j^0 \qquad u_i = \beta \sum_{j=1}^{|P|} \frac{w_{ij}}{\sqrt{d_i}\sqrt{d_j}} p_j + (1-\beta) u_i^0$$

α, β – hyper-parameter to weight the importance of the graph structure and prior query vector

$\alpha, \beta \in [0,1]$

equivalent matrix form

$$\mathbf{p} = \alpha S^T \mathbf{u} + (1-\alpha)\mathbf{p}^0 \qquad \mathbf{u} = \beta S \mathbf{p} + (1-\beta)\mathbf{u}^0 \qquad \boxed{\text{BiRank iteration}}$$

$S = D_u^{-\frac{1}{2}} W D_p^{-\frac{1}{2}}$ (symmetric normalization of weight matrix W)

BiRank:

Randomly initializes $\mathbf{p}$ and $\mathbf{u}$ ⇒ iteratively executes BiRank iteration until convergence

– time complexity of BiRank iteration

$O(|P| \cdot |U|)$ mainly due to $S^T\mathbf{u}$ and $S\mathbf{p}$

However, S typically very sparse

⇒ only need to account for non-zero entries

⇒ $O(c|E|)$

c – number of iteration executed to converge

10 iterations are usually enough for convergence