Graph Embeddings.

<u>Idea:</u> Map the vertices of G onto points in a metric space M. so that ~~nodes~~ two similar vertices in G map to two close points in M. Formally $\mathcal{E}: V(G) \to X$.

$M = (X, d)$

<u>Classical Embeddings:</u> $M = (\mathbb{R}^k, L_p)$ where L_p is the distance function $\left(\sum_{i=1}^{k} |x_i - y_i|^p\right)^{1/p} = L_p(x,y)$

Examples $p=1$. $\sum_i |x_i - y_i|$ (Manhattan or Hamming metric).

$p=2$ $\sqrt{\sum_{i=1}^{k}(x_i - y_i)^2}$ (Euclidean distance).

$p=\infty$ $\max_{i=1}^{k}(x_i - y_i)$

~~Structural Embedding~~s:

If v and w are <u>similar</u> then we want $\mathcal{E}(v)$ and $\mathcal{E}(w)$ to be close: $d(\mathcal{E}(v), \mathcal{E}(w))$ should be small.

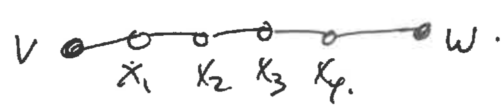
Proximity

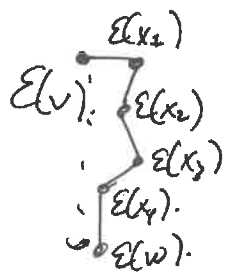
We need a way to measure how similar v and w are.
(similarity)

Proximity measure should be large when v and w are similar and small when v and w are not similar.

First-order proximity: $$S_1(v,w) = \begin{cases} 1 & \text{if } vw \in E(G) \\ 0 & \text{if } vw \notin E(G). \end{cases}$$

Intuition: If $\mathcal{E} : V(G) \to M$ ~~preserves~~ preserves first-order proximity then $d(\mathcal{E}(v), \mathcal{E}(w))$ should be related to shortest path in G from v to w.





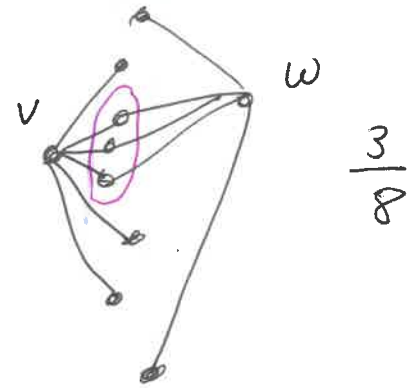
Second-Order Proximity:

$$s_2(v,w) = \frac{|N_G(v) \cap N_G(w)|}{\deg_G(v) \cdot \deg_G(w)} = \frac{\sum_{u \in V(G)} s_1(v,u) \cdot s_1(u,w)}{\sqrt{\sum_u s_1(v,u)} \cdot \sqrt{\sum_u s_1(u,w)}}$$

i^{th}-order proximity:

$$s_i(v,w) = \frac{\sum_u s_{i-1}(v,u) \cdot s_{i-1}(u,w)}{\sqrt{\sum_u s_{i-1}(u,v)} \cdot \sqrt{\sum_{s_{i-1}} (u,w)}}$$

Jaccard Proximity.



$$S_{J,1}(v,w) = \frac{|N(v) \cap N(w)|}{|N(v) \cup N(w)|}$$

$$S_{J,r}(v,w) = \frac{|N^r(v) \cap N^r(w)|}{|N^r(v) \cup N^r(w)|}$$

Proximity From Centrality Measures (Katz & Page Rank).

Recall centrality measures like Katz and PageRank.

$c(v)$ is ~~relative~~ a weighted sum of $\{c(w) : w \in N(v)\}$.

Katz$_\alpha$: $$c(v) = \alpha \cdot \sum_{w \in N(v)} c(w) + 1 \qquad (*)$$

Leads to matrix equation: $$\mathbf{c} = \underbrace{(I - \alpha A^T)^{-1}}_{n \times n} \mathbf{1}$$ 1-vector.

purpose of $\mathbf{1}$ is to implement the sum in $(*)$ for each $v \in V(G)$.

Before we do that: $\left((I - \alpha A^T)^{-1}\right)_{w,v}$ ~~represents the~~ represents

$\alpha \cdot c(w)$.

$$w \xrightarrow{\alpha \cdot c(w)} v$$

We can treat this as how much influence w has on v

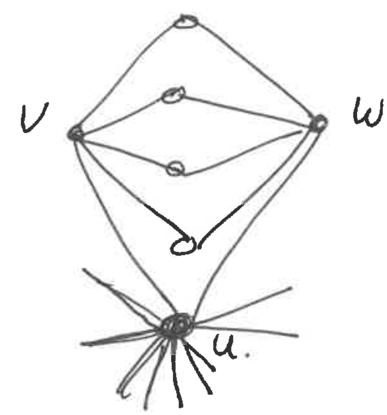
$$S_\alpha^{Katz} = \sum_{i=1}^{\infty} (\alpha \cdot A)^i = (I/\alpha - A)^{-1} \cdot A$$

Common Neighbours.

$$S^{CN} = A^2 \qquad s^{CN}(v,w) = |N(v) \cap N(w)|$$

Adamic-Adar

$$s^{AA}(v,w) = \begin{cases} 0 & \text{if } v = w. \\ \displaystyle\sum_{u \in N(v) \cap N(w)} \frac{1}{\ln(\deg(u))} & \end{cases}$$



Idea: If u has very high degree then the fact that u is adjacent to both v and w means less.

Linear-Algebra Embeddings Techniques.

To find embedding $\mathcal{E} : V(G) \to \mathbb{R}^k$

Local Linear Embedding.

adjacency matrix with entries in row x divided by $\deg(x)$.

Want: $\mathcal{E}(v)$ to be close to $\frac{1}{\deg^+(v)} \cdot \sum_{w \in N^+(v)} \mathcal{E}(w) = \sum_{w \in V(G)} \hat{a}(v,w) \cdot \mathcal{E}(w)$.

$\mathcal{E}$ can be described by a $k \times n$ matrix. E, where column v of E is $\mathcal{E}(v)$.

$$(I - \hat{A}) E^T$$

$I \cdot E^T = E \qquad \hat{A} E^T =$

difference between $\mathcal{E}(v)$ and $\sum_{w \in N^+(v)} \mathcal{E}(w)$.

$$E = K\,\begin{array}{c} \overbrace{v_1 \; v_2 \; \cdots \; n \qquad\qquad v_n} \\ \begin{bmatrix} \vdots & \vdots & \quad & \vdots \end{bmatrix} \end{array}$$

$$I \cdot E^T = \overset{I}{\begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & \ddots \end{bmatrix}} \overset{E^T}{\begin{bmatrix} \quad \end{bmatrix}} = \overset{E^T}{\begin{bmatrix} \quad \end{bmatrix}}$$

$$\hat{A} \cdot E^T = \overset{\hat{A}}{\begin{bmatrix} \frac{1}{4} & \\ \frac{1}{4} & \\ \frac{1}{4} & \\ \frac{1}{4} & \end{bmatrix}} \overset{E^T}{\begin{bmatrix} \quad \end{bmatrix}} = \begin{bmatrix} \quad \end{bmatrix}$$ ← average of neighbours.

$$(I - \hat{A})E^T = \begin{bmatrix} \bigcirc \\ \\ \\ \end{bmatrix}$$ ← difference between $E(v)$ and $\frac{1}{\deg(v)} \sum_{u \in N(v)} E(u)$.

Minimize $\|(I - \hat{A})E^T\|_F$ where $\|B\|_F = \sqrt{\left(\sum_i \sum_j b_{ij}^2\right)}$

↑

Euclidean length of vector.

How far from $\bar{0}$

Same as:

$$\|(I - \hat{A})E^T\|_F = \|((I - \hat{A}) \cdot E^T)^T\|_F = \|E(I - \hat{A})^T\|_F = \|E(I - \hat{A}^T)\|_F.$$

LLE: $\Phi(\mathcal{E}) = \sum_{v} \left(\mathcal{E}(v) - \frac{1}{\deg(v)} \sum_{w \in N^{+}(v)} \mathcal{E}(w) \right)^2$

$E\mathbf{1} = 0.$

$EE^T/n = I$

$$\Phi(E) = \| (I - \hat{A}) E^T \|_F = \mathrm{tr}\left(E (I-\hat{A})^T (I-\hat{A}) E^T \right)$$

Laplacian Eigenmaps.

$$\Phi(\mathcal{E}) = \sum_{v} \sum_{w \in N(v)} \| \mathcal{E}(v) - \mathcal{E}(w) \|^2 = \mathrm{tr}(E L E^T).$$

$$L = D - A$$

HOPE.

$$\Phi(\mathcal{E}) = \sum_{v} \sum_{w} \left(S(v,w) - \mathcal{E}(v) \cdot \mathcal{E}(w) \right)^2 = \| S - E^T E \|_F$$