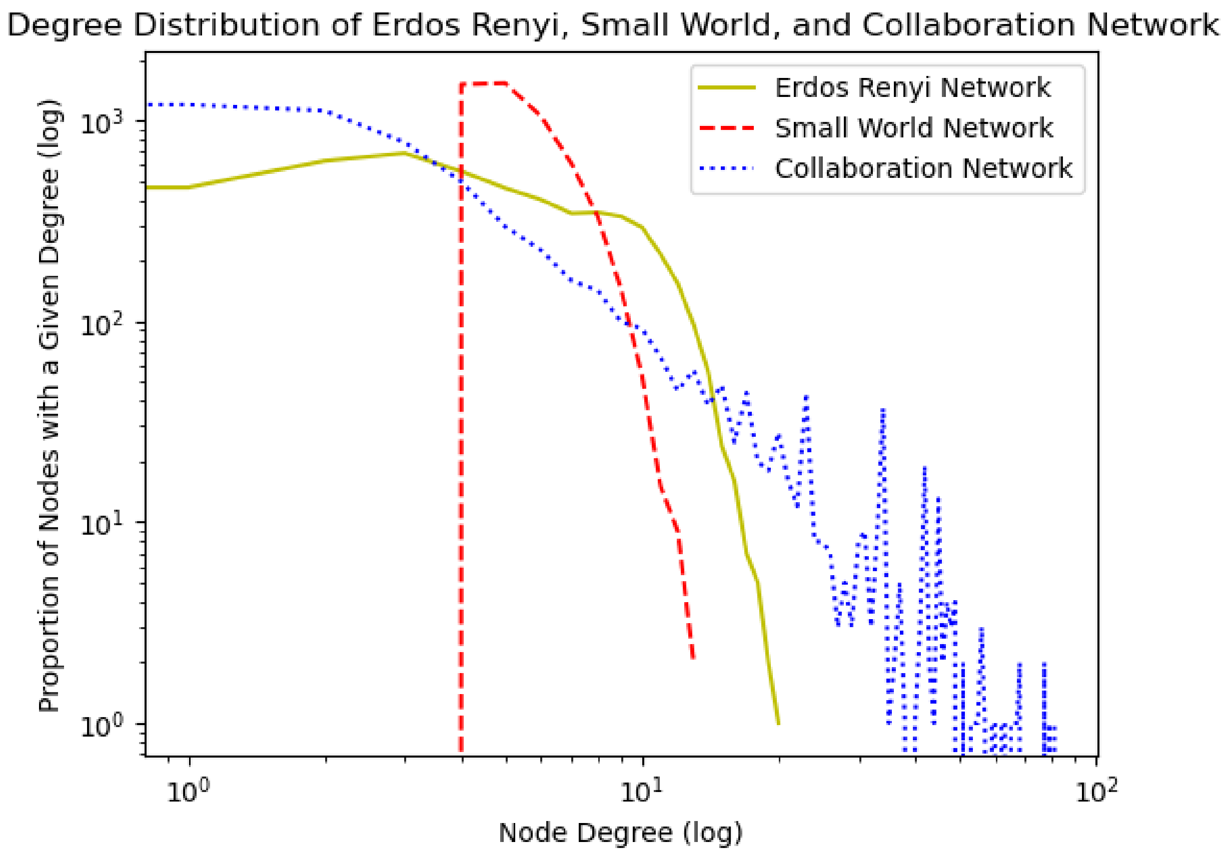
- The Small World Network is with narrow node degrees range
- Erdos Renyi Network is similar to the Collaboration Network at the low degree range but cannot cover the high degree range.
- The degree distribution of Collaboration Network is wider, with both more high-degree and low-degree nodes.

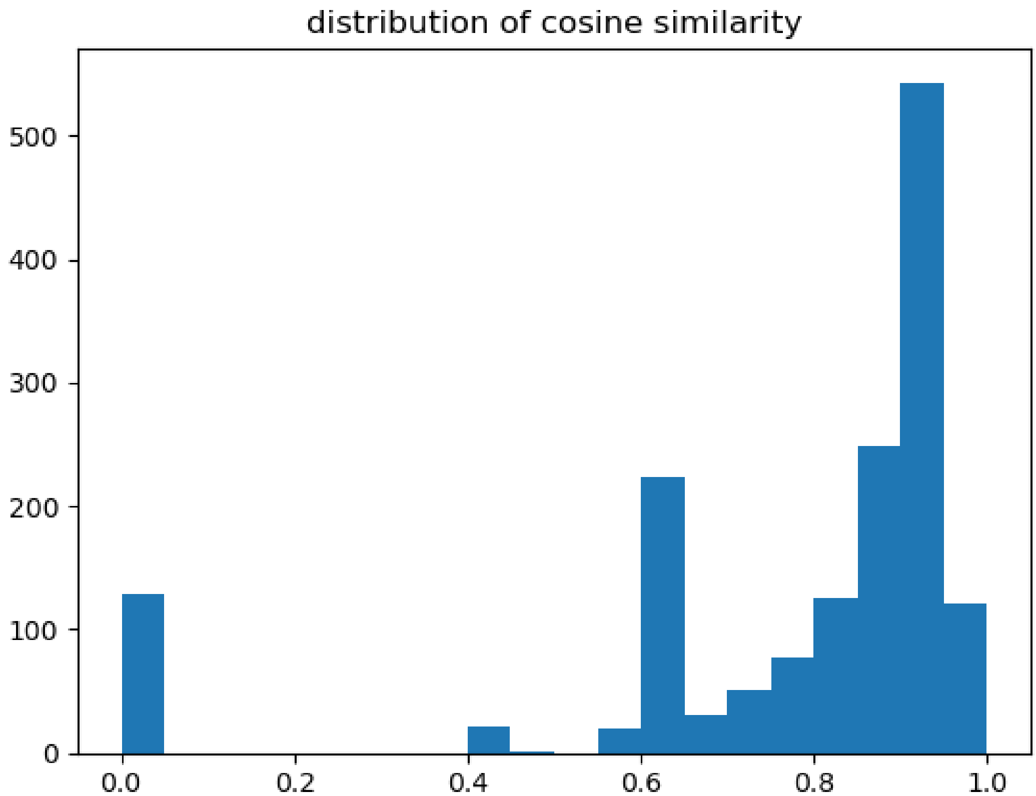


- Clustering Coefficient for Erdos Renyi Network: 0.001975
- Clustering Coefficient for Small World Network: 0.297949
- Clustering Coefficient for Collaboration Network: 0.529636

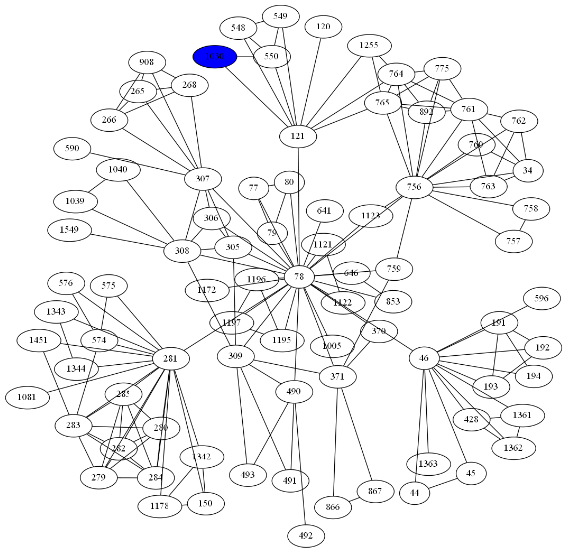
Collaboration Network is with largest average clustering coefficient. Authors may prefer to collaborate in their own network, resulting in a larger average clustering coefficient.

- Feature vector of node 9 is:[6 4 7]
- Top 5 nodes are:[(415, 0.9987327913808193), (286, 0.9970704162402443), (288, 0.9970704162402443), (16, 0.9950371902099892), (17, 0.9950371902099892)]

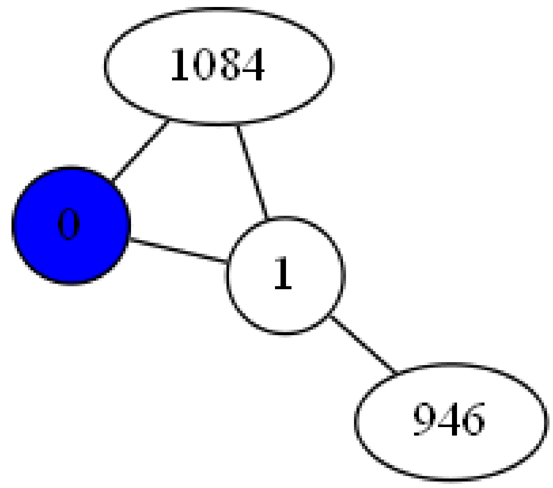
- Feature vector of node 9 is: [6. 4. 7. 2.5 1.83333333 6.5 15. 11. 39. 2.5 1.83333333 6.5 4.31944444 3.02777778 6.70833333 10.16666667 7.33333333 16.66666667 15. 11. 39. 25.91666667 18.16666667 40.25 61. 44. 100.]
- Top 5 nodes are: [(973, 0.9977857208153201), (415, 0.9967754809774014), (296, 0.9922005791814624), (189, 0.9908884045260696), (275, 0.9908884045260696)]



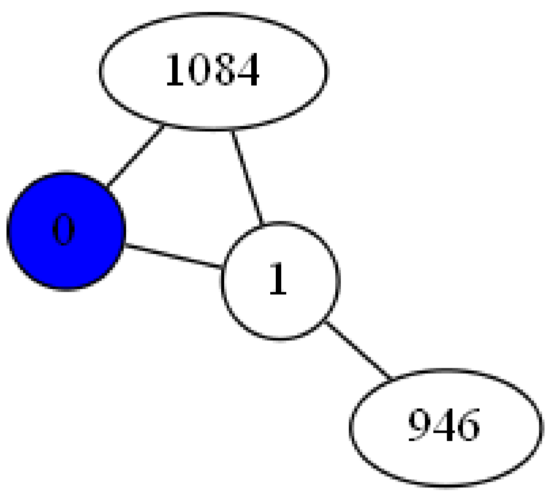
4 spikes can be observed. We have 4 groups: [0.0, 0.05], [0.4, 0.45], [0.6, 0.65], [0.9, 0.95]



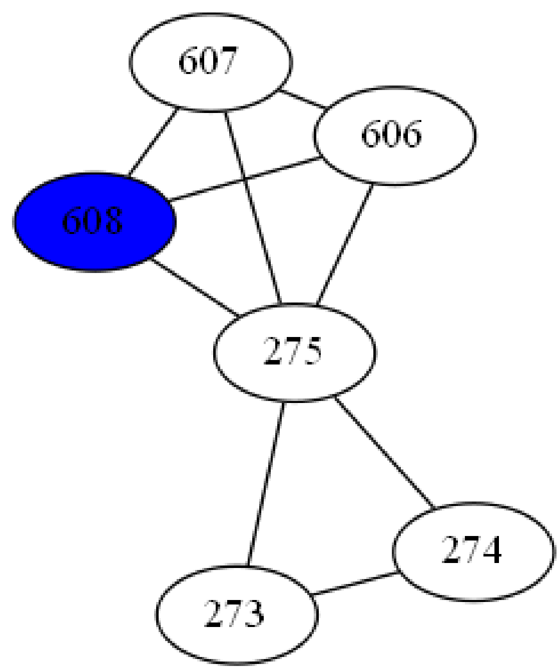
(a) $[0.0, 0.05]$



(b) $[0.4, 0.45]$

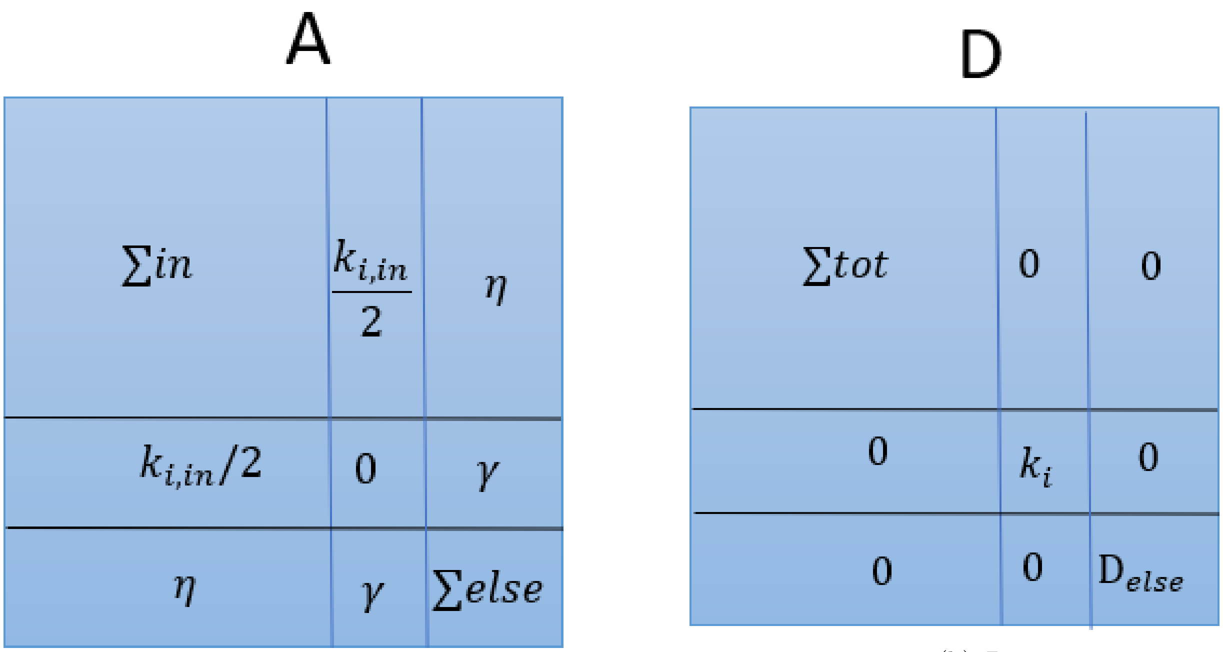


(c) $0.6, 0.65$



(d) $[0.9, 0.95]$

Figure 1: subgraphs



(a) A

(b) D

Figure 2: A and D

Node i is a single node with no self-loop. $\sum A$ for node i is 0, as shown in Fig 2. where

$$\gamma = k_i - \frac{k_{i,in}}{2}$$

$$\eta = D_{else} - \gamma - \sum else = D_{else} - k_i - \frac{k_{i,in}}{2} - \sum else$$

$$Q_{pre} = \frac{1}{2m}(\sum in - \frac{(\sum tot)^2}{2m} + 0 - \frac{(k_i)^2}{2m} + \sum else - \frac{(D_{else})^2}{2m}) \tag{1}$$

$$= \frac{\sum in + \sum else}{2m} - (\frac{\sum tot}{2m})^2 - (\frac{\sum k_i}{2m})^2 - (\frac{D_{else}}{2m})^2 \tag{2}$$

where Q_{pre} is the modularity of C and i before i joins in C.

$$Q_{post} = (\frac{(\sum in + k_{i,in})}{2m}) - (\frac{(\sum tot + k_i)^2}{2m}) + \frac{\sum else}{2m} - (\frac{D_{else}}{2m})^2 \tag{3}$$

where Q_pre is the modularity of C and i after i joins in C.

$$\Delta Q = Q_{pre} - Q_{post} \tag{4}$$

For graph H,

- the weight of any edge between two distinct nodes in H is 1
- the weight of any self-edge in H is 6
- the modularity of H is $Q = 4 * [12/5 - (14/56)^2] = 17/28 \approx 0.607$

For graph J,

- the weight of any edge between two distinct nodes in H is 2
- the weight of any self-edge in J is 13
- the modularity of j is $Q = 2 * [26/56 - (28/56)^2] = 3/7 \approx 0.429$

For graph H_{big},

- the weight of any edge between two distinct nodes in H_{big} is 1
- the weight of any self-edge in H_{big} is 6
- the modularity of H_{big} is $Q = 32 * [12/448 - (14/448)^2] = 185/224 \approx 0.826$

For graph j_{big},

- the weight of any edge between two distinct nodes in J_{big} is 2
- the weight of any self-edge in J_{big} is 13
- the modularity of J_{big} is $Q = 16 * [26/448 - (28/448)^2] = 97/112 \approx 0.866$

Modularity optimization algorithms have limitations in detecting small communities in large networks.Since in a graph, modularity will likely increase when we decrease the number of modularity. Considering the equation 9

$$Q = \frac{1}{2m}(-2cut(S) + \frac{1}{m}vol(S)vol(\bar{S})),$$

when the communities are fixed, merging several small (intuitive) communities will likely increase Q (cut decrease and vol*vol increase).

- (i) For $L = D - A$, consider adding an edge between node i and node j), the $(i, j)th$, $(j, i)th$ entries of A will change from 0 to 1 and the $(i, i)th$, $(j, j)th$ entries will both add 1. That means the $(i, i)th$, $(j, j)th$ entries of ΔL will increase 1, and the $(i, j)th$, $(j, i)th$ entries of ΔL will change from 0 to -1. ΔL exactly equals to $(e_i - e_j)(e_i - e_j)^T$.
- (ii) $x^{\mathrm{T}} Lx = \sum_{(i,j)\in E} x^{\mathrm{T}}(e_i - e_j)(e_i - e_j)^T x$. We know $x^{\mathrm{T}}(e_i - e_j) = (x_i - x_j)$ and $(e_i - e_j)(e_i - e_j)^T x = (x^{\mathrm{T}}(e_i - e_j))^T = (x_i - x_j)$. so $x^{\mathrm{T}} Lx = (x_i - x_j)^2$.
- (iii) If node i and node j in the same S, $x_i = x_j$, otherwise

$$\begin{aligned} x^{\mathrm{T}} Lx &= cut(S)(x_i^2 + x_j^2 - 2x_i x_j) \\ &= cut(S)(vol(\bar{S})/vol(S) + vol(S)/vol(\bar{S}) + 2) \\ &= cut(S)(vol(\bar{S}) + vol(S))^2/(vol(\bar{S})vol(S)). \end{aligned}$$

  Since $NCUT(S) = cut(S)(vol(\bar{S}) + vol(S))/(vol(\bar{S})vol(S))$, so $x^{\mathrm{T}} Lx = (vol(\bar{S}) + vol(S)) * NCUT(S) = C * NCUT(S)$. $C = 2m$

- (iv)

$$\begin{aligned} x^{\mathrm{T}} De &= \sum_i x_i D_{(i,i)} \\ &= \sum_{i\in S} \sqrt{\frac{vol(\bar{S})}{vol(S)}} D_{(i,i)} + \sum_{i\in \bar{S}} \sqrt{\frac{vol(S)}{vol(\bar{S})}} D_{(i,i)} \\ &= \sqrt{\frac{vol(\bar{S})}{vol(S)}} vol(S) + \sqrt{\frac{vol(S)}{vol(\bar{S})}} vol(\bar{S}) \\ &= 0 \end{aligned}$$

- (v)

$$\begin{aligned} x^{\mathrm{T}} Dx &= \sum_i x_i^2 D_{i,i} \\ &= \sum_{i\in S} \frac{vol(\bar{S})}{vol(S)} D_{i,i} + \sum_{i\in \bar{S}} \frac{vol(S)}{vol(\bar{S})} D_{i,i} \\ &= vol(\bar{S}) + vol(S) \\ &= 2m \end{aligned}$$

let $z = D^{1/2}x$, the optimization problem can be rewritten as

$$\begin{aligned}&\text{minimize}_{z \in \mathcal{R}^n} \frac{z^{\mathrm{T}} \tilde{L} z}{z^{\mathrm{T}} z} \\ &\text{s.t.} z^{\mathrm{T}} D^{1/2} e = 0, z^{\mathrm{T}} z = 2m,\end{aligned}$$

where $\tilde{L}$ is symmetric and normalized, the eigenvectors of $\tilde{L}$ are orthogonal and $0 \leq v_i$. Then z can be rewritten as the polynomial of these eigenvectors, i.e., $z = \sum_i a_i v_i$, where a_i are weights and v_i are the eigenvectors of $\tilde{L}$. We also have $v_i v_j = 0$ when $i \neq j$ and $v_i v_j = 1$ when $i = j$

Then, the optimization objective will be

$$\begin{aligned}\frac{z^{\mathrm{T}} \tilde{L} z}{z^{\mathrm{T}} z} &= \frac{(\sum_i a_i v_i)^{\mathrm{T}} \tilde{L} \sum_i a_i v_i}{2m} \\ &= \frac{\sum_i a_i^2 \lambda_i}{2m} \\ &(\sum_i a_i^2 = 2m)\end{aligned}$$

Notably, $z^{\mathrm{T}} D^{1/2} e = 0 = 0 * \mathrm{D}^{1/2}\mathrm{e}$ and $\tilde{L} D^{1/2} e = D^{-1/2} L e = D^{-1/2}(De - Ae) = 0 = 0 * \mathrm{D}^{1/2}\mathrm{e}$, i.e., $D^{1/2}e$ is the eigenvector v_1 of z and $\tilde{L}$ (with the corresponding eigenvalue to be 0-smallest eigenvalue). Assume $0 \leq \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$, then we have

$$\begin{aligned}z^{\mathrm{T}} v_1 &= a_1 v_1^{\mathrm{T}} v_1 \\ &= 0\end{aligned}$$

So $a_1 = 0$. To minimize the objective, we have $a_2 = \sqrt{2m}$ as the coefficient of the second smallest eigenvalue λ_2 and $a_i = 0$ $(i > 2)$. Then $x = (2m)^{1/2} D^{-1/2} v$

We have $cut(S) = \sum_{i \in S, j \in \bar{S}} A_{i,j}$ and $vol(S) = \sum_{i \in S} d_i$.

$$
\begin{aligned}
Q(y) &= \frac{1}{2m} \sum_{1 \leq i,j \leq n} [A_{i,j} - \frac{d_i d_j}{2m}] I_{yi=yj} \\
&= \frac{1}{2m}[2m - 2cut(S) - \frac{1}{2m}(vol(S)^2 + vol(\bar{S})^2)] \\
&= \frac{1}{2m}[2m - 2cut(S) - \frac{1}{2m}((vol(S) + vol(\bar{S}))^2 - 2vol(S)vol(\bar{S}))] \\
&= \frac{1}{2m}(-2cut(S) + \frac{1}{m} vol(S) vol(\bar{S}))
\end{aligned}
$$

Information sheet
CS224W: Machine Learning with Graphs

Assignment Submission Fill in and include this information sheet with each of your assignments. This page should be the last page of your submission. Assignments are due at 11:59pm and are always due on a Thursday. All students (SCPD and non-SCPD) must submit their homework via GradeScope (http://www.gradescope.com). Students can typeset or scan their homework. Make sure that you answer each (sub-)question on a separate page. That is, one answer per page regardless of the answer length. Students also need to upload their code on Gradescope. Put all the code for a single question into a single file and upload it.

Late Homework Policy Each student will have a total of *two* late periods. *Homework are due on Thursdays at 11:59pm PT and one late period expires on the following Monday at 11:59pm PT*. Only one late period may be used for an assignment. Any homework received after 11:59pm PT on the Monday following the homework due date will receive no credit. Once these late periods are exhausted, any assignments turned in late will receive no credit.

Honor Code We strongly encourage students to form study groups. Students may discuss and work on homework problems in groups. However, each student must write down their solutions independently, i.e., each student must understand the solution well enough in order to reconstruct it by him/herself. Students should clearly mention the names of all the other students who were part of their discussion group. Using code or solutions obtained from the web (GitHub/Google/previous year's solutions etc.) is considered an honor code violation. We check all the submissions for plagiarism. We take the honor code very seriously and expect students to do the same.

Your name: ______________________________

Email: ____________________ **SUID:** ____________________

Discussion Group: ______________________________

I acknowledge and accept the Honor Code.

(Signed) ______________________________