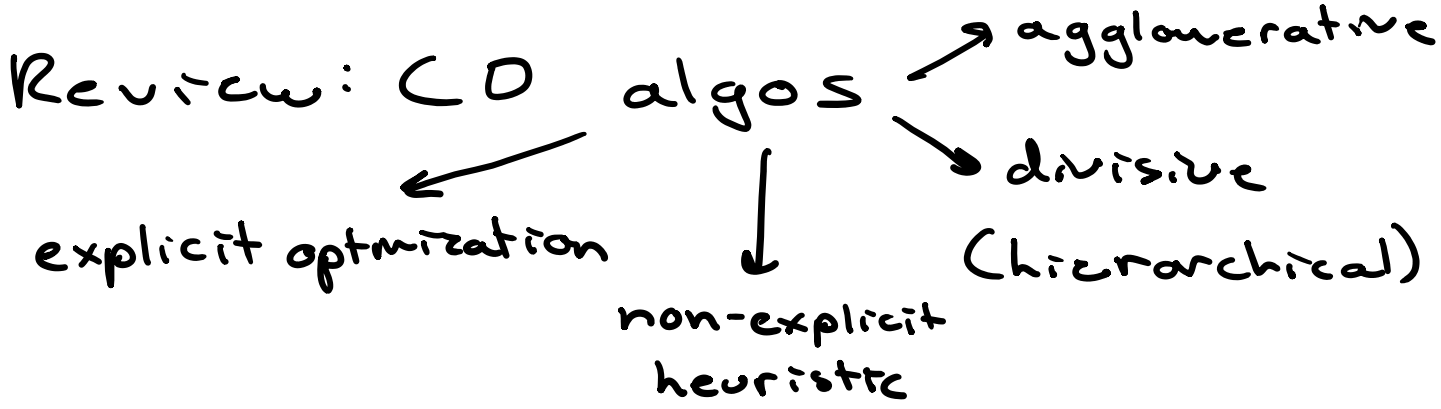
Lecture 13 - Community Detection Evaluation

Monday, March 16, 2026 9:00 AM



Optmization:

Modularity (Louvain)

Issues → resolution limit

→ simple graph probabilities

Edge Cut (Label Prop)

Issue → optmal is single community

Conductance

Issues → not a good global measure

→ tough to formulate as explicit optmization

Evaluation:

As observed, directly using the above measures can have several drawbacks

Other options:

- How many comms. output? (vs. expected)
- Community size distribution (vs. expected)

→ Issue: not always known

Gold Standard (for evaluation)

↳ comparing against a known "ground truth" solution

Issue: limited data available

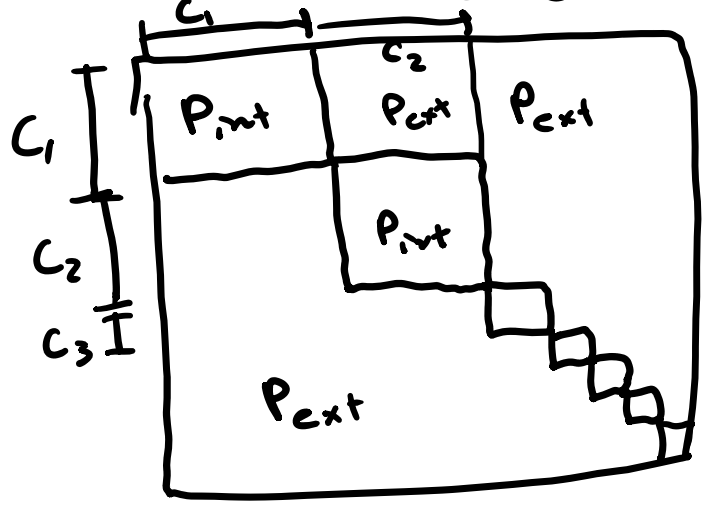
↳ companies keep their data hidden / private (API limited usefulness)

Solution: random graphs with a known community structure

Stochastic block model:

→ random graph

→ probabilities defined within blocks of the adjacency matrix



$P_{int} >> P_{ext}$

P_{int} = internal edge prob. in a community

P_{ext} = external edges between comms.

Random graph → G(n, p)

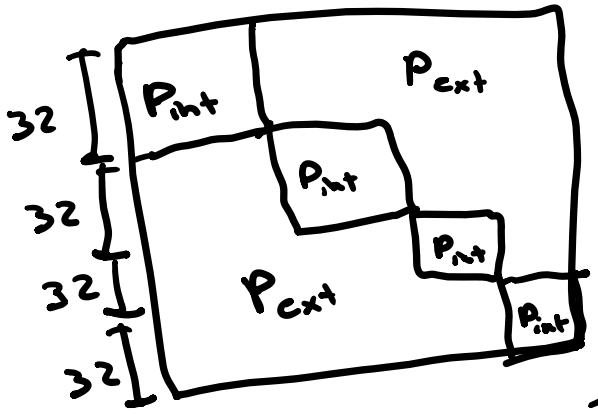
→ Erdös-Renyí

→ every edge between n vertices exists with probability p

Girvan-Newman (GN) Benchmark

SBM with $n = 128$

and 4 32 vertex communities



mixing parameter

↳ how defined our communities are

$$\mu_{GN} = \frac{k_{ext}}{k_{ext} + k_{int}}$$ ← avg. degree

$\mu_{GN} = 0 \rightarrow$ very defined comms.

$\mu_{GN} = 1 \rightarrow$ no defined comms.

GN $\frac{1}{3}$ SBMs in general

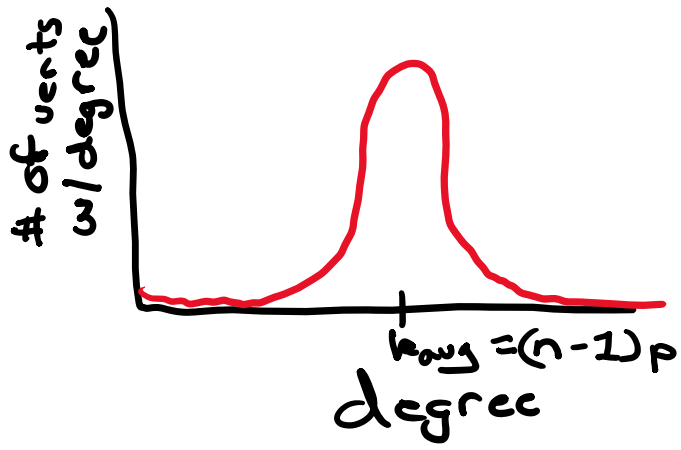
Good: Can control comm sizes + # (distribution)

Can control $|V(G)|, |E(G)|$

Can control comm. definition

↳ relative density (via μ)

Bad: $G(n,p)$ models aren't representative of real degree distribution



degree distribution of a G(n,p) graph

LFR Benchmark

↳ main goal: real degree distributions
varying comm. size distributions

the benchmark:

→ number of vertices

→ avg. degree

→ min/max comm size

→ exponents for degree and community size distributions

→ $\mu = \frac{|E_{ext}|}{|E_{ext}| + |E_{int}|}$ ← number of external edges; ← # of internal edges

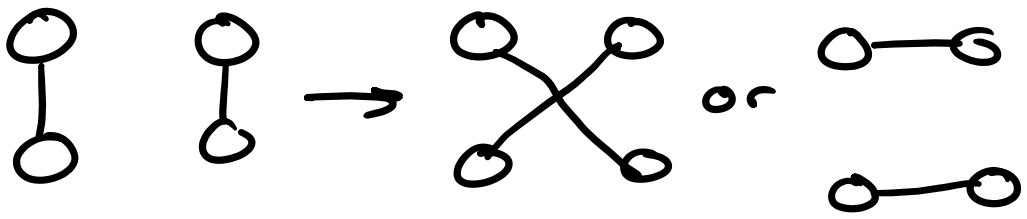
Algo: construct distributions

↳ assign degrees → verts

↳ construct edges (configuration model)

↳ verts → comms.

↳ rewire edges to hit μ



(degrees are maintained, but μ is updated)

Benchmark study

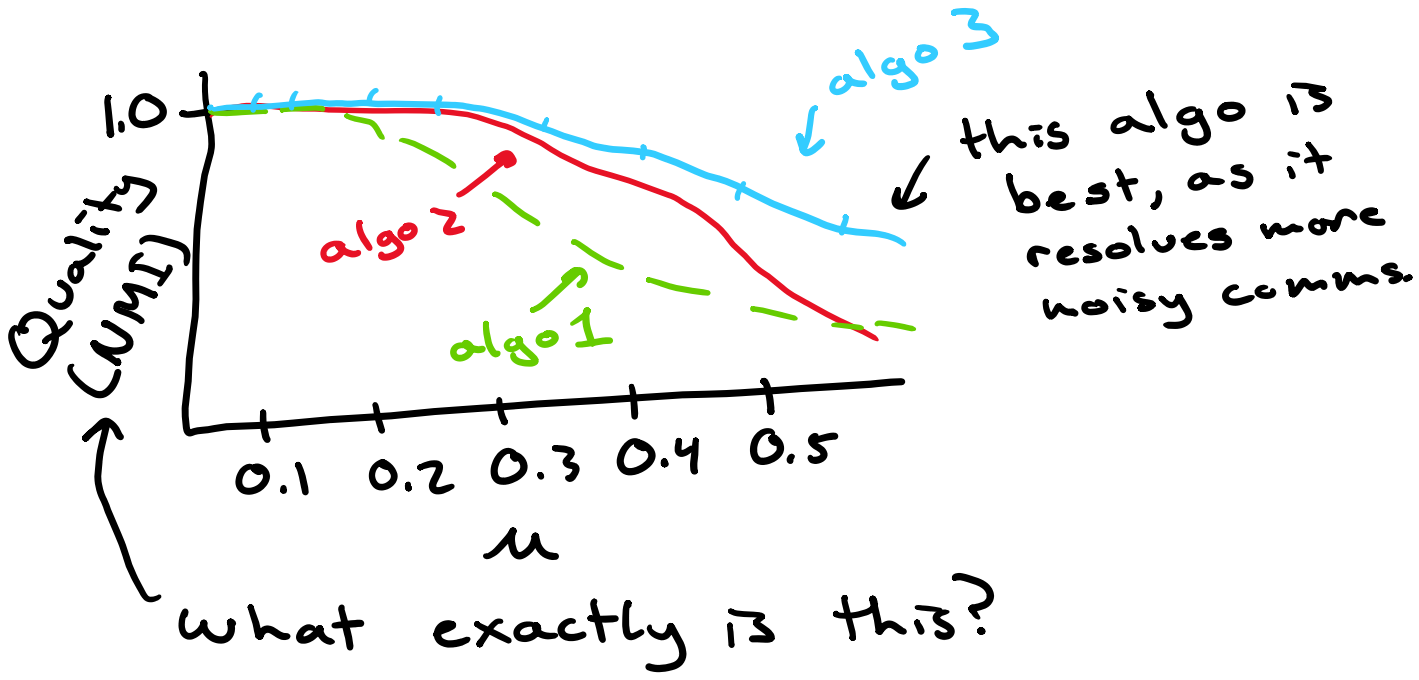
→ run algos to compare performance

(comm. detection algos on LFR graphs)

→ vary μ to see when algos start to fail

e.g.) $\mu = [0.1, 0.2, 0.3, \ldots 0.9]$

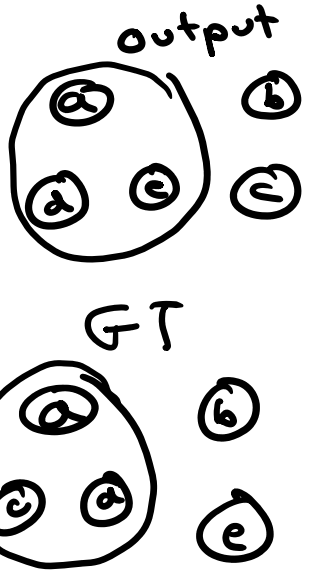
What we get:



Q: Given same output communities, how can we objectively compute same measure of quality relative to a known ground truth?

A: NMI (normalized mutual information)

vertex	comm. output	GT
a	1	3
b	2	1
c	3	3
d	1	3
e	1	4



e 1 7 

Contingency table

↳ enumerating all overlapping pairs in $|U| \times |V|$ table

$|U| = R$ = outputs (number of comms)

$|V| = C$ = GTs

$$T = \begin{array}{c|cccc} & \overbrace{1} & \overbrace{2} & \overbrace{3} & \overbrace{4} \\ \hline 1 & 0 & 0 & 2 & 1 \\ R\ 2 & 1 & 0 & 0 & 0 \\ 3 & 0 & 0 & 1 & 0 \end{array}$$

2 vertices labeled '1' per U '3' per V

$$T_{ij} = |U_i \cap V_j|$$

$$NMI = \frac{MI}{\frac{1}{2}(H(U) + H(V))}$$

← mutual info

← entropy of U, V

To compute:

$$H(U) = \sum_{i}^{R} P_U(i) \log(P_U(i))$$

← how widely variable the label distribution

$$P_U(i) = \frac{|U_i|}{\quad} = \text{prob. that a vertex}$$

$$P_U(i) = \frac{|U_i|}{N}$$ = prob. that a vertex is randomly in U_i comm

$H(V)$, $P_V(j)$ are computed similarly

$$P_{UV}(i,j) = \frac{T_{ij}}{N}$$ ← prob. that a vertex is in both U_i & V_j

$$MI = \sum_{i}^{R} \sum_{j}^{C} P_{UV}(i,j) \log\left(\frac{P_{UV}(i,j)}{P_U(i)\,P_V(j)}\right)$$

↑ measure of the overlap in comm. assignments (i.e., how close does knowing one inform you of the other)

$$NMI = \frac{MI}{\frac{1}{2}(H(U)+H(V))}$$

↑ the amount of overlap relative to how random the label distributions are

(i.e., the overall significance of the overlap)