Next Monday two things due

- HW #2
- Peer grade HW #1

Network measures for collections of nodes: Cohesion

- density, clustering, connectivity, flow, partioning

Density

Edge density = $\frac{|E|}{\binom{N_v}{2}}$, How often does occur? is another type of density

SNA uses a triad census to understand

Motifs ≃ subgraphs that may occur more frequently

Clustering (transitivity)

Fraction of that close to form — Closing triads

- related to homophily - neighbors of neighbors are likely to be neighbors

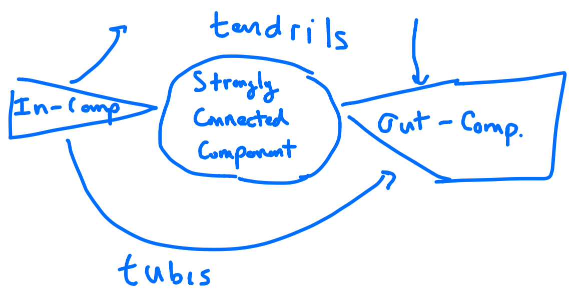
Connectivity

Q: does the graph seperate into different components? If not how close?

Component decomposition

- giant component (typically all that is used to model) $O(N_v)$
- smaller component $O(\log N_v)$
- isolates $O(1)$

Components in Digraphs



Graph Partioning

Goal: Partition G into subsets that demonstrate a "cohesiveness" w.r.t. to the graph

Rmk: NP-complete type problems

Cohesive

Subsets of vertices that

(a) well connected

(b) seperated from the remaining

Find a partion of $V = \{C_1, .., C_K\}$ s.t. $|E(C_k)| \gg |E(C_k, C_{k'})| \ \forall k \neq k'$.

Hierarchial Clustering

- Greedily optimizes some loss function – locally optimal solutions

- either agglomerative successive coarsening through merging

  divisive successive refinement through splitting

Choice of cost measure: (i) single linkage (ii) common linkage

Euclidean Dist $x_{ij} = \sqrt{(A_{ik} - A_{jk})^2}$

Neighborhood

$$x_{ij} = \frac{|N_{v_i} \Delta N_{v_j}|}{d_{N_v} + d_{v_{v-1}}}$$

Modularity f_{ij} = fraction of edges in G connecting C_i with C_j

$$\mathrm{mod}(C) = \sum_{k=1}^{K} [f_{kk}(C) - f_{kk}^*]^2$$

L model specific parameter