

## Midterm Review

- Path: a sequence of nodes that each consecutive pair is connected by an edge.
  - Component: set of nodes that can be traversed through a path are part of the same component.  
Note: a single node by itself IS a component.
  - Gatekeeper: a node that if it is removed, the graph is broken into two or more components.
  - Bridge: an edge that if it is removed, the graph is broken into two or more components.
  - Local Bridge: an edge that if it is removed, there would still exist an alternative path between the two nodes that were connected through that edge. For example if edge AB is a local bridge, then if edge AB is removed, there would still exist another path between AB with the length of more than two, that connects A to B.
  - Triadic Closure: A satisfies the triadic closure property if AB and AC are both strong edges and BC is either weak or strong.
  - Clustering co-efficient of node A : number of the edges that currently exist between node A's neighbors / total number of the edges that can exist between A's neighbor.
  - Embeddedness of an edge in a network is the number of common neighbors the two endpoints have.  
a bridge has embeddedness of zero.
  - **Neighborhood overlap of node A and B: number of shared neighbors between A and B / total number of the neighbors of both node A and B**
    - o **PLEASE PLEASE DO NOT DOUBLE COUNT!**
- ⊆ Density of a component: number of the present edges in the component / number of all possible edges.
- Number of all possible edges in a graph with n nodes is  $(n * (n-1))/2$
  - Test for homophily: if the proportion of cross edges to total number of present edges in the graph is significantly less than  $2pq$  then there is evidence for homophily.
    - o example: coloring a graph with two colors then p and q for the colors.
    - o p: probability of being a male
    - o q: probability of being a female
    - o  $qp + pq =$  probability of edge connecting a male to a female
    - o
    - o p: 1/3 q: 2/3 the  $2pq = 4/9$
    - o
    - o then we check the graph to see how many cross edges out of total present edges exist. Then we compare that with the  $2pq$  value. If that value was smaller than  $2pq$  then there exist a homophily in the graph.
  - pivot is a node that lies on every shortest path between two nodes.
  - local gatekeeper is a node that is connected two other nodes that there exists no direct edge between those two nodes. A is connected to B, A is connected to C but C is not connected to B. Therefore A is a local gatekeeper.
  - Balanced triangle: A triangle that has either 1 or 3 positive edges.
-