Social and Information Networks

Module 1 - GraphsMatrices

Reference Book:

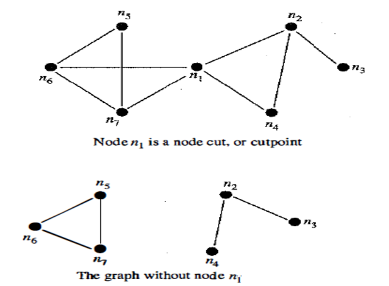
Wasserman Stanley, and Katherine Faust. (2009). Social Network Analysis: Methods and Applications, Structural Analysis in the Social Sciences.

Graph

Connectivity of Graphs:

- Connectivity of a graph is a function of whether a graph remains connected when nodes and/or lines are deleted.
- Cutpoints: A node, n_i, is a cutpoint if the number of components in the graph that contains n_i, is fewer than the number of components in the subgraph that results upon deleting n_i, from the graph.
  - Consider graph G with node set N, which includes node n_i, and the subgraph G_s, with node set $N_s = N - n_i$, that results from dropping n_i and all of its incident lines from graph G. Node n_i is a cutpoint if the number of components in G is less than the number of components in G_s

Graph -Connectivity



Node n_1 is a node cut, or cutpoint

The graph without node n_1

Example of a cutpoint in a graph

Graph -Connectivity

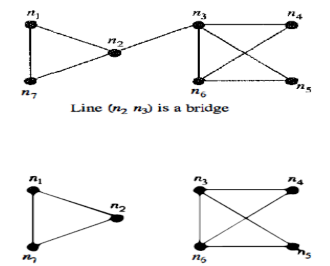
- Cutpoint: Example: In a communications network, an actor who is a cutpoint is critical, if that actor is removed from the network, the remaining network has two subsets of actors, between whom no communication can travel.
- Concept of a cutpoint can be extended from a single node to a set of nodes
  - If the set is of size k, then it is called a k-node cut.
  - A cutpoint is a 1-node cut
- Cutset: A set of nodes necessary to maintain the connectedness of a graph

Graph -Connectivity

Bridges:

- A line that is critical to the connectedness of the graph
- A bridge is a line such that the graph containing the line has fewer components than the subgraph that is obtained after the line is removed (nodes incident with the line remain in the subgraph).
- Removal of a bridge leaves more components than when the bridge is included

Graph -Connectivity



Line $(n_2\ n_3)$ is a bridge

Graph without line $(n_2\ n_3)$

Example of a bridge in a graph

Graph -Connectivity

- Consider graph G with line set L, with line l_k as the bridge and the subgraph G_s, with line set $L_s = L - l_k$, that results from removing l_k from graph G. Line l_k is a line cut if the number of components in G is less than the number of components in G_s
- A bridge is a 1-line cut.
- An l-line cut is a set of l lines that, if deleted, disconnects the graph.
- In social networks, a bridge is a critical tie, or a critical interaction between two actors.
- Connectivity of a graph is one measure of its "cohesiveness" or robustness.

Graph -Connectivity

Two measures of the connectivity of a graph:

- Node-Connectivity
- Line-Connectivity

A graph is cohesive if, for example, there are

- relatively frequent lines
- many nodes with relatively large degrees
- relatively short or numerous paths between pairs of nodes

- Cohesive graphs have many short geodesics and small diameters, relative to their sizes
- If a graph is not cohesive then it is "vulnerable" to the removal of a few nodes or lines
- A vulnerable graph is more likely to become disconnected if a few nodes or lines are removed

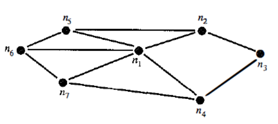
Graph -Connectivity

Node (Point)-Connectivity:

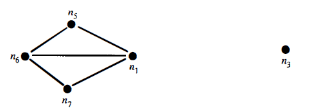
- Point-connectivity or node-connectivity of a graph, K(G), is the minimum number K for which the graph has a K-node cut
- K is the minimum number of nodes that must be removed to make the graph disconnected
- If the graph is disconnected, then K = 0, since no node must be removed.
- If the graph contains a cutpoint, then K = 1 (Removal of the single node leaves the graph disconnected)
- If a graph contains no node whose removal would disconnect the graph, but it contains a pair of nodes whose removal together would disconnect the graph, then K = 2, since two is the minimum number of nodes that must be removed to make the graph disconnected.
- Higher values of K indicate higher levels of connectivity of the graph

Graph –Connectivity

Example: 2-node cut



n_2 and n_4 comprise a 2-node cut



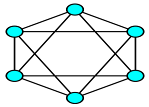
The graph without n_2 and n_4

Connectivity in a graph

Graph -Connectivity

Node (Point)-Connectivity :

- Removing any number of nodes less than K does not make the graph disconnected
- For any value **k less than K**, the graph is said to be **k-node connected**
- A complete graph has no cutpoint
  - All nodes are adjacent to all others
  - Removal of any one node would still leave the graph connected
  - To disconnect a complete graph, remove g – 1 nodes, resulting in a trivial graph (g = 1)
- Example:



Graph -Connectivity

Line-Connectivity :

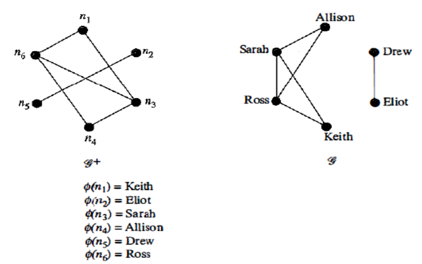
- Line-connectivity or edge-connectivity of a graph, λ(G), is the minimum number λ for which the graph has a λ -line cut.
- λ is the minimum number of lines that must be removed to disconnect the graph or leave a trivial graph
- If λ(G) >= l, the graph is said to be l-line connected, since l is the minimum number of lines that must be removed to make the graph disconnected
- The Larger the node-connectivity or the line-connectivity of a graph is, the less vulnerable the graph is to become disconnected

Graph

Isomorphic Graphs :

- Two graphs, G and G*, are isomorphic if there is a one-to-one mapping from the nodes of G to the nodes of G* that preserves the adjacency of nodes.
- Each node in G is mapped to one (and only one) node in G*, and each node in G* is mapped to one (and only one) node in G.
- Mapping preserves adjacency
  - if nodes that are adjacent in G are mapped to nodes that are adjacent in G*, and
  - nodes that are not adjacent in G* are mapped to nodes that are not adjacent in G and vice versa
- Two isomorphic graphs are identical on all graph theoretic properties (same number of nodes, the same number of lines, the same diameter , …)

Graph



Isomorphic graphs

Graph

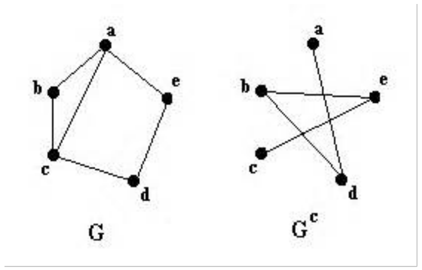
Special kinds of Graphs:

- Complement
- Tree
- Bipartite Graph

Complement:

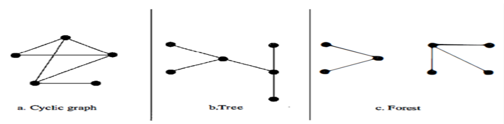
- Complement, G^c, of a graph, G, has the same set of nodes as G
  - A line is present between an unordered pair of nodes in G^c if the unordered pair is not in the set of lines in G
  - A line is not present in G^c if it is present in G
    - If nodes n_i and n_j are adjacent in G , then n_i and n_j are not adjacent in G^c
- Line sets for these two graphs have no intersection at all, and their union is the set of all possible lines (all unordered pairs of nodes)

Graph
Example: Complement



Graph

- Tree: A graph that is connected and is acyclic
- Forest: A graph that is disconnected (has more than one component) and contains no cycles
- Number of lines in a tree or forest equals the number of nodes minus the number of components of the graph



Cyclic and acyclic graphs

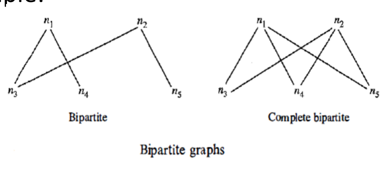
Graph

Bipartite graph:

- A graph is bipartite
  - If the nodes in a graph can be partitioned into two subsets, N1 and N2, so that every line in L is an unordered pair of nodes in which one node is in N1 and the other node is in N2
- There are two subsets of nodes
- All lines are between nodes belonging to different subsets
- Nodes in a given subset are adjacent to nodes from the other subset, but no node is adjacent to any node in its own subset

Graph

- Example:



- Two subset of nodes are N1 and N2
- N1 = {n_1, n_2 } and N2 = {n_3 , n_4, n_5}

Graph

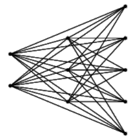
Complete bipartite graph

- Every node in N1 is adjacent to every node in N2
- Denoted $K_{g1,g2}$, where g1 is the number of nodes in N1, and g2 is the number of nodes in N2
- Example:
  - A two-mode network with two sets of actors and a relation linking actors in one set to actors in the second set can be represented as a bipartite graph

Graph

s-partite graph :

- Partitioning of the nodes into 's' subsets so that all lines are between a node in N_i and a node in N_j, where i != j.
- All lines are between nodes in different subsets and no nodes in the same subset are adjacent
- A graph is a complete s-partite graph if all pairs of nodes belonging to different subsets are adjacent.
- No line is incident with two nodes belonging to the same subset.
- Example: Complete s-Partite graph



Graph

Example Bipartite Graph : A social network described by a bipartite graph is the set of monetary donations transacted between corporations in a specific geographic area, and the non-profit organizations headquartered in this area.

- Initially place all firms, both corporations and nonprofit organizations, into a single actor set, N
- Measure the flows of donations among these firms. Since the non-profit organizations usually have limited cash resources and thus can not support themselves financially, they must rely on the corporations for donations
- The only lines in this graph connect corporations to non-profit organizations
- Hence , we have a bipartite graph, with the corporations residing in set N1 and non-profit organizations in set N2

Directed Graph

- A relation is directional if the ties are oriented from one actor to another.
- Example 1: The import/export of goods between nations is an example of a directional relation.
  - In a social network representing trade among nations, the ties are directional and the graph representing such ties must be directed
- Example 2: Choices of friendships among children. The claim of friendship is directed from one child to another

Directed Graph

- A directional relation can be represented by a directed graph, or digraph
- Consists of a set of nodes representing the actors in a network, and a set of arcs directed between pairs of nodes representing directed ties between actors.
- A Directed Graph G_d(N, L) consists of set of nodes N = {n_1, n_2, n_3, ..., n_g }, and a set of arcs, L = {l_1, l_2, l3, ..., l_L }.
- Each arc is an ordered pair of distinct nodes, l_k = < n_i , n_j >
- There are g(g - 1) possible arcs in L

Directed Graph

Subgraphs – Dyads

- Consisting of two nodes and the possible arcs between them
- There are four possible states for each dyad
  - Null dyad
  - Asymmetric dyad
  - Mutual or Reciprocal dyad

Directed Graph

Nodal Indegree and Outdegree

- Indegree of a node, $d_I(n_i)$, is the number of nodes that are adjacent to n_i
  - Number of arcs of the form l_k = <n_j ,n_i>
- Outdegree of a node, $d_O(n_i)$, is the number of nodes that are adjacent from n_i
  - Number of arcs of the form l_k = <n_i ,n_j>
- Outdegrees are measures of expansiveness and indegrees are measures of receptivity, or popularity

Directed Graph

- Mean Indegree and Mean Outdegree

$$\bar{d}_I = \frac{\sum_{i=1}^{g} d_I(n_i)}{g}$$

$$\bar{d}_O = \frac{\sum_{i=1}^{g} d_O(n_i)}{g}$$

- Since $\sum_{i=1}^{g} d_I(n_i) = \sum_{i=1}^{g} d_O(n_i) = L$
  - we have $\bar{d}_I = \bar{d}_O = \frac{L}{g}$

Directed Graph

- Variability : Quantifies how unequal the actors in a network are with respect to initiating or receiving ties
- Variance of the indegrees, $S_{D_I}^2 = \frac{\sum_{i=1}^{g}(d_I(n_i) - \bar{d}_I)^2}{g}$
- Variance of the outdegrees , $S_{D_O}^2 = \frac{\sum_{i=1}^{g}(d_O(n_i) - \bar{d}_O)^2}{g}$
- Density: proportion of arcs present in the graph $\Delta = \frac{L}{g(g-1)}$

Directed Graph

- Types of Nodes in a Directed Graph
  - *Isolate* if $d_I(n_i) = d_O(n_i) = 0$,
  - *Transmitter* if $d_I(n_i) = 0$ and $d_O(n_i) > 0$,
  - *Receiver* if $d_I(n_i) > 0$ and $d_O(n_i) = 0$,
  - *Carrier* or *ordinary* if $d_I(n_i) > 0$ and $d_O(n_i) > 0$
- Carrier node has both indegree and outdegree precisely equal to 1
- Ordinary node has indegree and/or outdegree greater than 1