Network Analysis

1 Structure

Density of a Graph   Defined as the percentage of non-zero edges.

$$D(G) = \frac{m}{\binom{n}{2}} = \frac{2m}{n(n-1)}$$

Random Graph Model $G(n,p)$ is a graph with n vertices, and for each pair v_i, v_j there is an edge with probability p.
If we want an average degree of k we choose $p = \frac{k}{n-1}$

Degree Sequence   The vertices of a graph ordered in descending order of their degree.

Durfee Square   A partition of n integers has Durfee number s if s is the largest number such that the partition contains at least s parts with value $\geq s$.
Example of a Durfee square of size 3:
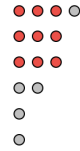
Dunbar's Number   is a suggested cognitive limit to the number of people with whom one can maintain stable social relationships. He proposed this number to be 150.

Scale-Free Network   A network is called scale-free, if the degree distribution of its graph $G = (V, E)$ (roughly) obeys a power-law.

Preferential-Attachment Graph   This leads to a Scale-Free Network. When creating a preferential-attachment graph we add the vertices one after another and add edges to a set of k existing vertices. The probability that the new vertex is connected to an existing vertex is proportional to that vertex' degree.

1.1 Core-Periphery

Split Graph   A Graph G is a split graph if there exists a partition $V = C \uplus P$ such that C is a clique and P an independent set (i.e. $G[C]$ is complete and $G[P]$ consists of only isolated vertices).
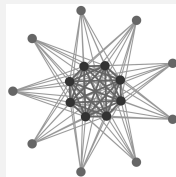
Splittance   is the split edit distance. This means it is the minimum number of edge additions and deletions needed to turn a given graph into a split graph.

Neighborhood Inclusion   Vertex v is dominated by w, $v \preceq w$, if $N(v) \subseteq N[w]$.

Threshold Graph   A undirected graph $G = (V, E)$ is a threshold graph if the neighborhood-inclusion preorder $\preceq$ is total, i.e. if $v \preceq w$ or $w \preceq v$ (or both) for all $v, w \in V$.

1.2 Small World

Characteristics of a Small World

1. Small Distances
2. Low Clustering ≤ 0.75 (social groups are assumed to be dense)
3. Bounded degree (as in Dunbar's number)

Clustering coefficient   Density of the neighbourhood of v

$$cc(v) = D(G[N(v)]) = \frac{\#Edges\ in\ N(v)}{\binom{|N(v)|}{2}}$$

Cluster Coefficient of G is the average over all vertices

$$cc(G) = \frac{1}{n} \cdot \sum_{v \in V} cc(v)$$

Wieners Index   Total of all Path lengths

$$W(G) = \sum_{s,t \in V} dist(s,t)$$

(k, d, r) - torus

k: Number of Nodes

d: Dimension of Nodes

r: Connected to all neighbours that are r steps away

Lemma   If $G = (V, E)$ is a (n, 1, r) - torus with $r \leq \frac{n-1}{4}$, then

1. $<deg> = 2r$
2. $<cc> = \frac{3(r-1)}{4(r-0.5)}$
3. $<dist> \geq \lfloor \frac{n-1}{4r} \rfloor$

1.3 Triad-Census

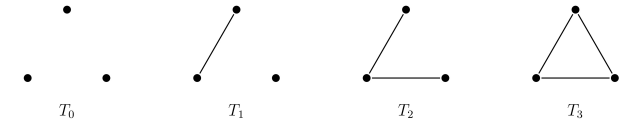This is a vector of length 4 that contains the number of occurrences of the different kinds of triangles in a graph:

Figure 2.2: The four non-isomorphic graphs on three vertices

2 Centrality

Let $G = (V, E)$ be a graph.

Degree Centrality   For a simple undirected graph $G = (V, E)$, degree centrality for a node $i \in V$ is defined as

$$c_D(i) = \deg(i)$$

2.1 Closeness Centrality

The independence of a point is determined by its closeness to all other points in the graph.

Closeness   is defined as

$$c_C(i) = \sum_{t \in V} dist(i,t)$$

Closeness Centrality   is defined as

$$c_C(i) = \left[\sum_{t \in V} dist(i,t) \right]^{-1}$$

General Closeness Centrality   is defined as

$$c_C(i) = \frac{R(i)}{n-1} \cdot \left[\frac{\sum_{t \neq i: i \to *t} dist(i,t)}{R(i)} \right]^{-1}$$

for all $i \in V$, where R(i) is the number of vertices reachable from i. The sum is over all nodes that are reachable over a directed path.

Harmonic closeness

$$c_H(i) = \frac{1}{n-1} \cdot \sum_{t \neq i : i \to *t} \frac{1}{dist(i,t)}$$

Radiality

$$c_R(i) = \frac{1}{n-1} \cdot \sum_{t \neq i : i \to *t} \frac{1}{diam(G) + 1 - dist(i,t)}$$

Farness

$$c_F(s) = \sum_{t \in V} dist(s,t) - 1$$

2.2 Betweenness Centrality

Betweenness is an index of the potential of a point to control the communication.

Dependency   Let $G = (V, E; \lambda)$ be a cost-valued graph. The (shortest-path) dependency of a pair of vertices $s \neq t \in V$ on another vertex $b \in V \setminus \{s,t\}$ is defined as

$$\delta(s,b,t) = \frac{\sigma(s,t|b)}{\sigma(s,t)} = \frac{\#shortest\ s-t\ paths\ over\ b}{\#shortest\ s-t\ paths}$$

where $\sigma(s,t)$ is the number of shortest (s,t) - paths and $\sigma(s,t|b)$ is the number of shortest (s, t) - paths with b as an inner vertex. We let $\delta(s,b,s) = 0 = \delta(s,s,t) = \delta(s,t,t)$.

$$\sigma(s,t|b) = \begin{cases} \delta(s,b) \cdot \delta(b,t) & if\ dist(s,t) = dist(s,b) + dist(b,t) \\ 0 & otherwise \end{cases}$$

> **Betweenness Centrality**   is defined for a vertex $i \in V$ as
>
> $$c_B(i) = \sum_{s \neq t \in V} \delta(s,i,t) = \sum_{s \neq t \in V} \frac{\sigma(s,t|i)}{\sigma(s,t)}$$
>
> This is the fraction of shortest $(s-t)$-paths with intermediary i.
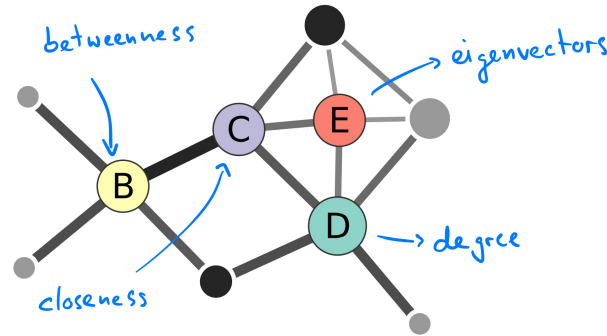
Dyadic dependency   of sender s on broker b

$$\delta(s,b) = \sum_t \delta(s,b,t) = \sum_{s \neq t \in V} \frac{\sigma(s,t|b)}{\sigma(s,t)}$$

Theorem   Total betweenness equals total closeness.

$$\sum_{b \in V} c_B(b) = W(G) - R(G) = \sum_{s \in V} c_F(s)$$

The total degree to which brokers b control shortest-path relations equals the total dependency of vertices s on intermediaries.

For each approach a different vertex can be central:

2.3 Generalization

2.3.1 Substitutions

Identity
$S(\preceq) = \{id\}$ yields $i \preceq j \iff N(i) \subseteq N(j)$

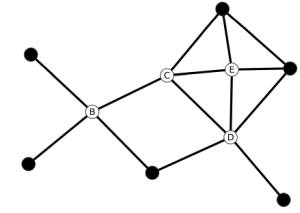**Transpositions**
$S(\preceq) = T \cup \{id\}$ yields $i \preceq j \iff N(i) \subseteq N[j]$
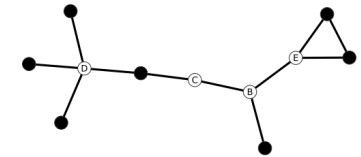(i having j as a neighbour can be compensated by j being it's own neighbour)
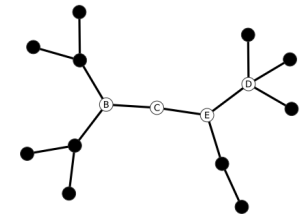
Everyone can replace anyone

$$S(\preceq) = S_A \text{ yields } i \preceq j \iff deg(i) \leq deg(j)$$

(a) order-minimal ($n = 10$, $m = 14$)

(b) size-minimal ($n = 11$, $m = 11$)

(c) smallest tree ($n = 15$, $m = 14$; not the only one)

Figure 4.4: Smallest graphs with distinct maximally central vertices according to betweenness (B), closeness and eccentricity (C), degree (D), and eigenvector (E) centrality. (Determined by Jan Hildenbrand listing candidate graphs using the nauty program.)