SC435
Introduction to Complex Networks-3

Centrality

- Many networks are heterogeneous $\Rightarrow$ wide variability in the properties and role of their elements.
- **Importance** of a **node** or a **link** in the network is measured by its position with respect to others.

Degree Centrality

- degree of a node (number of nearest neighbors)

$$C_D(i) = k_i = \sum_j A_{ij}$$

- normalized degree centrality

$$C_D^* = \frac{1}{n-1} C_D(i)$$

does not contain any information about the structure of the graph

Centrality

Closeness centrality

Measures how close an actor is to to all other actors in the network.

- Mean geodesic distance from i to j

$$l_i = \frac{1}{n}\sum_j d_{ij} = \frac{1}{n-1}\sum_{j(\neq i)} d_{ij}$$

  l_i gives a low value for more central vertices and high value for less central vertices.

- Closeness Centrality

$$C_C(i) = \frac{1}{\sum_j d_{ij}}$$

- Normalized closeness centrality

$$C_C^*(i) = \frac{1}{l_i} = \frac{n-1}{\sum_{j(\neq i)} d_{ij}}$$

Actors in the center can quickly interact with others (minimal number of steps to reach others)

Value's dynamic range is relatively close$\Rightarrow$ fluctuates as network changes.

Centrality

Betweenness centrality

Many phenomena taking place in networks are based on diffusion process, example disease spread, information spread in social networks, transport etc $\Rightarrow$ node is more central if it is more often involved in these processes.

- Betweeness centrality measures the extent to which a vertex lies on the shortest paths between vertices.
- Betweenness Centrality

$$C_B(i) = \sum_{s,t} \frac{\sigma_{st}(i)}{\sigma_{st}}$$

- Normalized Betweenness centrality

$$C_B^* = \frac{2}{(n-1)(n-2)} C_B(i)$$

Centrality

Eigenvector centrality

- Degree centrality assumes all neighbors are equal $\Rightarrow$ only the number of neighbors matter.
- In many cases importance of the node depends on the importance of its neighbors.
- How can we construct a centrality based on this idea?

Centrality

Eigenvector centrality (Bonacich (1987))

- Assume initially that everyone has a score $x_i = 1$ we update its centrality using the sum/scores of their neighbors

$$x_i^{'} = \sum_j A_{ij} x_j \Rightarrow \boxed{\mathbf{x}^{'} = A\mathbf{x}}$$

- After t steps we have a vector of centralities $\mathbf{x}(t)$ given by

$$\mathbf{x}(t) = A^t \mathbf{x}(0)$$

- Expressing $\mathbf{x}(0)$ as a linear combination of the eigenvalues of A, we have $\mathbf{x}(0) = \sum_i c_i \mathbf{v}_i$

$$\mathbf{x}(t) = A^t \sum_i c_i \mathbf{v}_i = \lambda_1^t \sum_i \left(\frac{\lambda_i}{\lambda_1}^t\right) \mathbf{v}_i$$

- $\mathbf{x}(t)$ converges to a limit given by the eigenvector corresponding to the maximum eigenvalue.

$$\lim_{t \to \infty} \mathbf{x}(t) \to c_1 \mathbf{v}_1 \lambda_1^t \qquad \boxed{A\mathbf{v} = \lambda \mathbf{v}}$$

Centrality

Eigenvector centrality

- Eigenvector centrality can be large either because (either/or or both)
  - a vertex has many neighbors
  - it has important neighbors
- Eigenvector centrality of all vertices are non-negative.
- Directed graphs:
  - Two sets of eigenvectors, since A is asymmetric.
  - in-degree (a vertex is important if it is referred to by others)

$$x_i = \frac{1}{\kappa_i} \sum_j A_{ij} x_j$$

Centrality

Katz centrality

- Problems with eigenvector centraility
  - Only vertices that are in a strongly connected component of two or more vertices, or the out-component of such a component can have non-zero eigenvector centrality.
  - Acyclic directed graph
- **Solution**: Add a constant (importance) to each vertex

$$x_i = \alpha \sum_j A_{ij} x_j + \beta$$

- **Katz Centrality:** With $\beta = 1$

$$\mathbf{x} = (I - \alpha A)^{-1}\mathbf{1} = \sum_n \alpha^n A^n$$

Centrality

Page Rank

- One problem with eigenvector centrality is that a page gets a high centrality since it has a in-degree from a high centrality page.
- The fix is to divide the centrality contribution of a vertex by its out-degree.
- Page rank centrality

$$x_i = \alpha \sum_j A_{ij} \frac{x_j}{k_j^{out}}$$

- Matrix Notation:

$$\mathbf{x} = \alpha A D^{-1} \mathbf{x} + \beta \mathbf{1}$$

- If there are any vertex with out-degree 0 then the first term becomes indeterminate $\Rightarrow$ artificially set k_j^{out} to 1.

Centrality

The random walk (surfer) view

- Page rank can be thought of as a model of user behavior.
- Assume a randim surfer (walker), who starts at a page and moves along the directed path by choosing links at random.
- Once in a while the user gets bored and starts on another random page.
- The steady state distribution, probability that the random surfer visits a page is its page rank score.
- $$P(v_i) = \frac{1-d}{N} + d \sum_{j \in \mathcal{N}(i)} \frac{P(v_j)}{k_j^{out}}$$

Centrality

Hubs and Authority

- **Authority:** Nodes that contain useful information on a topic of interest.
- **Hubs:** Nodes that point to authorities or tell us where the best cuthority can be found (points the search engine in the right direction).
- **HITS:** Hyperlink induced text search assigns two different centralities to any vertex
  - Authority centrality x_i
  - Hub centrality y_i

Centrality

HITS (Jon Kleinberg (Cornell))

- The authority centrality of vertex i is proportional to the sum of the hub centralities of the vertices that point to it.
- Hub centrality of a vertex is proportional to the sum of the authority centrality of the vertices that it points to.
- Authority centrality

$$x_i = \alpha_i \sum_j A_{ij} y_j$$

$$\mathbf{x} = \alpha \, A\mathbf{y}$$

- Hub Centrality

$$y_i = \beta \sum_j A_{ji} x_j$$

$$\mathbf{y} = \beta A^T \mathbf{x}$$

Centrality

HITS

- The authority and hub centralities are the eigenvectors of AA^T and A^TA with the same leading eigenvalue

$$AA^T\mathbf{x} = \lambda\mathbf{x}$$

$$A^TA\mathbf{y} = \lambda\mathbf{y}$$

- All the eigenvalues are the same for AA^T and A^TA.
- $AA^T \Rightarrow$ Bibliography coupling
- $A^TA \Rightarrow$ Cocitation coupling

swap

- λ is the dominant eigenvalue (Perron-Frobenius Theorem)
- vertices not cited by any can have zero authority centrality but still have non zero hub centraility.