CS4423: Networks

Week 6, Part 2: Centrality Measures

Dr Niall Madden

School of Maths, University of Galway

Week 6 (19+20 Feb 2025)

These slides include material by Angela Carnevale.

Outline

Today's notes are split between these slides, and a Jupyter Notebook.

Slides are at:

https://www.niallmadden.ie/2425-CS4423

What is it that makes a node in a network important?

Key nodes in networks can be identified through **centrality measures**: a way of assigning "scores" to nodes that represents their "importance". However, what it means to be central depends on the context.

Examples

- In a friendship network, who is most popular?
- In a epidemiology network, who is most likely to get infected?
- In a banking, which institution poses the greatest danger to the system should it fail?

Accordingly, in the context of network analysis, a variety of different centrality measures have been developed.

Centrality Measures

Measures of centrality include:

- **Degree Centrality** which is just the degree of the node. It can be important in e.g., transport networks.
- **Eigenvector Centrality**, defined in terms of properties of the network's **adjacency matrix**.
- **Closeness Centrality**, defined in terms of a nodes **distance** to other nodes on the network.
- **Betweenness Centrality**, defined in terms of **shortest paths**.

Degree Centrality

Definition (Degree Centrality)

In a (simple) graph $G = (X, E)$, with $X = \{0, \ldots, n-1\}$ and adjacency matrix $A = (a_{ij})$, the degree centrality c_i^D of node $i \in X$ is defined as

$$c_i^D = k_i = \sum_j a_{ij},$$

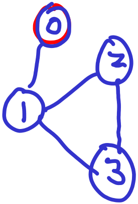
where k_i is the degree of node i.

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

ROW SUMS

$$\begin{pmatrix} 1 \\ 3 \\ 2 \\ 2 \end{pmatrix}$$

Example:



$$c_0^D = 1$$
$$c_1^D = 3$$
$$c_2^D = 2$$
$$c_3^D = 2.$$

In some cases, this measure can be misleading, since it depends–among other things–on the order of the graph. A better measure is then the following.

Normalized Degree Centrality

The **normalized degree centrality** C_i^D of node $i \in X$ is defined as

$$C_i^D = \frac{k_i}{n-1} = \frac{c_i^D}{n-1} \left(= \frac{\text{degree centrality of node } i}{\text{number of potential neighbors of } i} \right).$$

Note: in a directed graph one distinguishes between the in-degree and the out-degree of a node and defines the in-degree centrality and the out-degree centrality accordingly.

We now recall from important facts from **Linear Algebra**.

Eigenvalues and Eigenvectors

Let A be a square $n \times n$ matrix. An n-dimensional vector, $\boldsymbol{v}$, is called an **eigenvector** of A if

$$A\mathrm{v} = \lambda \mathrm{v}$$

for some scalar (number), λ, which is called an **eigenvalue** of A.

Example:

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}.$$

so $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ is an eigenvector associated with eigenvalue $\lambda = 2$.

Finished here Wed.

- When is a real-valued matrix, one usually finds that λ and $\boldsymbol{v}$ are *complex valued*. However, if A is symmetric then they are *real valued*.
- A may have up to n eigenvalues: λ_1, λ_2, ..., λ_n.
- The **spectral radius** of A is $\rho(A) := \max\{|\lambda_1|, |\lambda_2|, \dots, |\lambda_n|\}$
- If v is an eigenvector associated with the eigenvalue λ, so too is any non-zero multiple of v

Centrality

The basic idea of eigenvector centrality is that **a node's ranking in a network should relate to the rankings of the nodes it is connected to**.

More specifically, up to some scalar λ, the centrality c_i^E of node i should be equal to the sum of the centralities c_j^E of its neighbouring nodes j.

In terms of the adjacency matrix $A = (a_{ij})$, this relationship is expressed as

$$\lambda c_i^E = \sum_j a_{ij} c_j^E,$$

which in turn, in matrix language is

$$\lambda c^E = A c^E,$$

for the vector $c^E = (c_i^E)$, which then is an eigenvector of A.

So c^E is an eigenvector of A! (But which one???)

How to find c^E and/or λ?

If the network is small, one could use the usual method (although it is almost never a good idea)

1. Find the *characteristic polynomial* $p_A(x)$ of A, as *determinant* of the matrix $xI - A$, where I is the $n \times n$ identity matrix);
2. Find the *roots* λ of $p_A(x)$ (i.e. scalars λ such that $p_A(\lambda) = 0$);
3. Find a *nontrivial solution* of the linear system $(\lambda I - A)c = 0$ (where 0 stands for an all-0 column vector, and $c = (c_1, \dots, c_n)$ is a column of *unknowns*).

For large networks, there are much better algorithms, such as the **Power Method** that we'll study later (in the Week 6 – Part 3 Jupyter Notebook).

Presently, we'll learn that the adjacency matrix always has one eigenvalue which is greater than all the others.

First, some definitions:

Irreducible Matrix

A matrix A is called **reducible** if, for some simultaneous permutation of its rows and columns, it has the block form

$$A = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}.$$

A is **irreducible** if it is not reducible.

Important: The adjacency matrix of a simple graph G is irreducible if and only if G is connected.

Non-negative matrix

A matrix $A = (a_{ij})$ is **non-negative** if

$$a_{ij} \geq 0 \qquad \text{for all } i, j.$$

For simplicity, we usually write $A \geq 0$.

Important: Adjacency matrices are examples non-negative matrices.

There are similar concepts of, say, positive matrices (nothing to do with positive definite!!), negative matrices.

Of particular importance are **positive vectors**: $v = (v_i)$ is positive if $v_i > 0$ for all i. We write $v > 0$.

The Theorem

Theorem (Perron-Frobenius Theorem 1907/1912)

Suppose that A is a square, nonnegative, **irreducible** *matrix. Then:*

- *A has a real eigenvalue $\lambda = \rho(A)$ and $\lambda > |\lambda'|$ for any other eigenvalue λ' of A. λ is called the* **Perron root** *of A*
- *λ is a simple root of the characteristic polynomial of A (so has just one corresponding eigenvector)*
- *There is an eigenvector, $\boldsymbol{v}$ associated with λ, such that $v > 0$.*

For us this means:

- The adjacency matrix of a connected graph has an eigenvalue that positive, and greater in magnitude than any other.
- It has an eigenvector, $\boldsymbol{v}$ that is positive.
- v_i is the Eigenvector Centrality node i.