

Network of Thrones

A Song of Math and Westeros

A Primer on Network Analysis

The Networks Around Us

We have always been aware of the powerful characteristics of the networks that we inhabit, even if we couldn't fully grasp the forces behind them. A chance encounter at a party reveals a common friend who lives thousands of miles away. Folklore tells us that the best way to find a new job is to "network" via acquaintances and friends-of-friends. All of us have surfed the web, suddenly realizing that just a few clicked links have taken us worlds away from our original intentions.



A complex network by [Martin Grandjean \(http://www.martingrandjean.ch/\)](http://www.martingrandjean.ch/) via Wikimedia Commons

We live in a highly interconnected world, at many different scales. Connections are what networks are about: two nodes are joined by an edge when they are related in a specified way. We are tied to our friends. Cities are connected by roads and airline routes. Flora and fauna are bound together in a food web. Countries are involved in trading relationships. But not all networks are physical: the World Wide Web is a virtual network of information.

Until recently, thinking about networks at grand scales was intractable. Different fields reacted to this deficiency in their own ways. Sociologists limited themselves to small data sets (fewer than 50 people). Mathematicians focussed on purely aesthetic models unencumbered by applications. Scientists used simplified models that ignored the underlying networks. Our modern Information Age has produced a wealth of data about the complex networks that tie us together. In response, the field of Network Science has arisen, drawing from mathematics, computer science, sociology, economics, and the sciences.

The last two decades have witnessed the development of a coherent methodology for analyzing, understanding and interpreting network data. This specialty is evolving before our eyes, and is sure to undergo further evolution in the years to come. Network Science is highly interdisciplinary: the field synthesizes multiple traditions to grapple with understanding the structure of networks, and how this structure effects the dynamic interactions between its nodes.

A Crash Course on Empirical Network Analysis

On this site, we will apply some of the tools from Network Science to everyone's favorite complex system: the kingdoms of Westeros and Essos. We perform an exploratory analysis of networks of characters, focussing on two main questions:

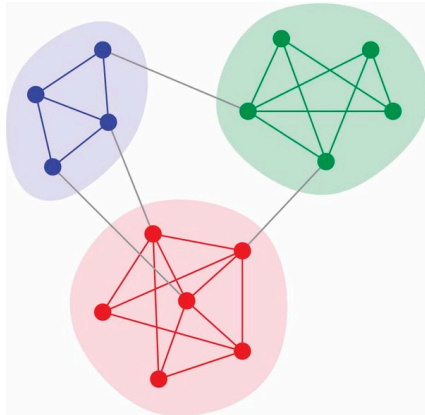
- Which characters naturally belong together, forming coherent communities within the network?
- Which characters play an important role in the network?

These are actually complicated questions. We explore them both using mathematical tools that capture their essence.

Communities

The definition of a "community" is a slippery one. Intuitively speaking, a community is a subset of the network that forms a self-contained and coherent sub-network. One common metric that people use for community detection is that there should be lots of edges within communities, and fewer edges between communities. This idea is captured by a quantity called **modularity** that I won't define explicitly here. Given a network, we use standard techniques to split the network into communities so that modularity is maximized (approximately, anyway). One nice feature of this process is that we discover the number of communities, rather than picking the optimal number of communities up front.

Once we have partitioned a network into communities, we color the nodes according to their community, and choose a network layout that emphasizes the community structure. For example, here is a network with three communities:

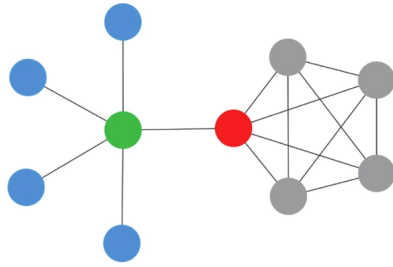


A network with three communities, image from <https://data.graphstream-project.org>

Centralities

Not all nodes are created equal. Some nodes play an outsized role in the network, either by having many connections, or by being strategically positioned to help connect distant parts of the network. In fact, there are many ways that a node can be important or influential. In our analysis, we focus on five different centrality measures, each of which captures a different dynamic. These centrality measures have been used in diverse fields, from sociology to economics to computer science. There is no single "right" centrality measure. As with other Data Science techniques, it is essential to interpret these quantities with respect to our expert knowledge in the underlying domain.

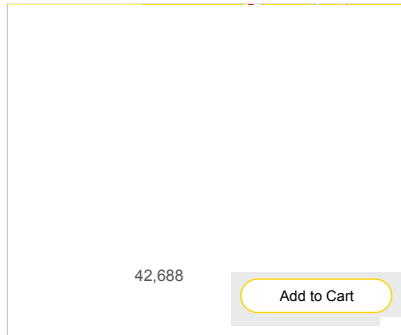
For example, consider the following network. Which vertices are important? Which vertex is the most important?



Network centrality example from [Wikipedia's Centrality page](https://en.wikipedia.org/wiki/Centrality) (<https://en.wikipedia.org/wiki/Centrality>).

It's pretty clear that the green and red vertices are the most important ones. This importance reflects more than just connection count (the gray vertices have nearly as many neighbors). The red and green vertices are essential to the connectivity of the network. But which is more important: green or red? It really depends; what kind of network is this? If this is a trading network, then green has a lot of power, since the blue vertices can only trade with green. But if this is a political network, then red might be more powerful, since red can promote the consensus opinion among a close-knit group.

Advertisements



REPORT THIS AD

We will be looking at a network whose nodes are characters. Two nodes are connected by an edge when those two characters interact (in some way or another). These edges are weighted by the total number of interactions. Characters who interact frequently will be connected by a high weight edge. Acquaintances will be connected by a low weight edge. Complete strangers will not be connected by any edge whatsoever.

Here are the five different centrality measures that we will use.

Degree Centrality

The degree centrality of a node is the number of other nodes that are directly connected to it via an edge. This is just a raw count of the number of people that the character interacted with at least once.

Weighted Degree Centrality

The weighted degree centrality is the sum of the weights of the edges incident with the node. This is the total number of interactions involving the character.

Eigenvector Centrality

This is weighted degree centrality with a feedback loop. Having connections to "important" people makes you more important as well. In this measure, you get full credit for knowing someone important, even if you don't know them very well. This measures how powerful your network is (in theory), regardless of whether you are using your network to its fullest potential.

