Network Explorer

Gauden Galea

14 Jun 2014

Network Explorer is a simple app built as an exercise for the Developing Data Products course on Coursera. It is intended primarily as a teaching tool for exploration of the connectivity patterns in two distinct types of networks.

Basic Concepts and Definitions

These two diagrams illustrate two different networks, each composed of 20 *nodes*, or *vertices*. Each connection between two nodes is called an *edge*. The number of neighbours connected to node N is called the *degree* of node N.
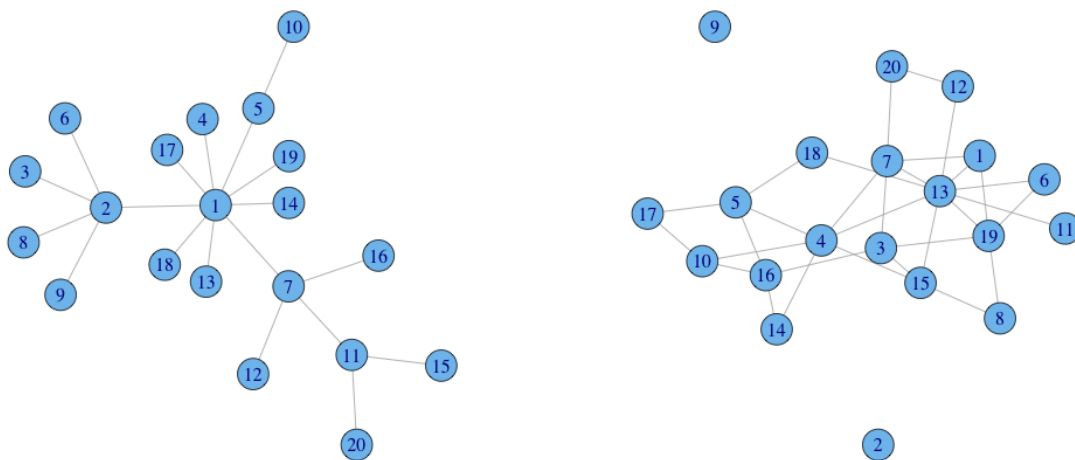
The two networks have been created in different ways:

1. In the graph on the left, the more connected a node is, the more likely it is to receive a connection from any new node joining the network. This creates a pattern of connections characterised by a few highly connected nodes (a few nodes of high degree), while most of the other nodes have low degree (a very small number of neighbours).
2. In the graph on the right, all edges have an equal probability of existing, i.e. any two nodes have a known probability of being connected by a common edge. This results in a pattern of more evenly distributed degree across the nodes without any nodes standing out as being highly connected.

Using the App

Network Explorer is a simple app visualising these two styles of graph creation. It allows interactive exploration of the connectivity patterns that are generated in the two families of network. There are three choices to be made:

1. In the left hand panel, one first selects the graph type: preferential attachment or random attachment. These correspond to the left and right graph styles int he diagram above.
2. One next selects the number of nodes using a slider widget. The number of nodes is limited to between 5 and 100, allowing sufficient range to explore patterns without creating network diagrams that are coimpletely indecipherable.
3. If one has selected the random attachment style, another slide appears where one can set the probability of attachment. This is the overall probability that any two nodes in the network will be connected.

Any choices made in the left hand panel result immediately in a change in the output on the right hand panel. This panel is divided into two tabs, entitled "Network" and "Degree Distribution":

- The **Network** tab displays a ball-and-stick diagram of the resulting network. It also displays a few numerical indices: the *average path length* is the average number of edges that lie between nodes in connected clusters, the *number of clusters* is the number of nodes in each group of connected nodes, and the *graph density* is the ratio of the number of actual edges to the number of all possible edges between nodes.
- The **Degree Distribution** tab displays a relative distribution chart showing the proportion of nodes having a given number of neighbours.

Exploration

1. Explore the patterns of graph that result from selecting different graph types while keeping the number of nodes constant.
2. For a given number of nodes, explore the graph density differences between the two graph types
3. Selecting the random attachment graph type, explore the patterns that emerge when varying the probability while keeping a constant number of nodes or vice versa. For example, for a given probability of attachment, how many are needed before all the nodes are connected in one large group (the number of clusters becomes 1)
4. Explore the effect of all these on the degree distribution. Which is the

Credits

- Network Explorer is a Shiny app
- built with R and Rstudio.
- The igraph package is used for the network creation and analysis.
- The ggplot2 poackage is used for plotting distributions