Random Graphs

G. Huber

18Jul2016 / Final Project Presentation

Outline

Outline

1. Introduction
   - Introduction
2. Random Graph Models
   - Erdös - Rényai model
   - Barabási - Albert model
   - Watts - Strogatz model
3. RandomGraph application
4. Summary

Random Graph
A definition

Using the terminology in probability, a random graph is a random variable defined in a probability space with a probability distribution.[Chung2016]

In layman's terms, we first put all graphs on n vertices in a lottery box and then the graph we pick out of the box is a random graph. (In this case, all graphs are chosen with equal probability.) [Chung2016]

Random Graphs
Uses

- Random graphs have been used to model the growth of the internet.
- Random graphs have been used to help analyze social networks.
- Random graphs have been used to help model epidemiology

Outline

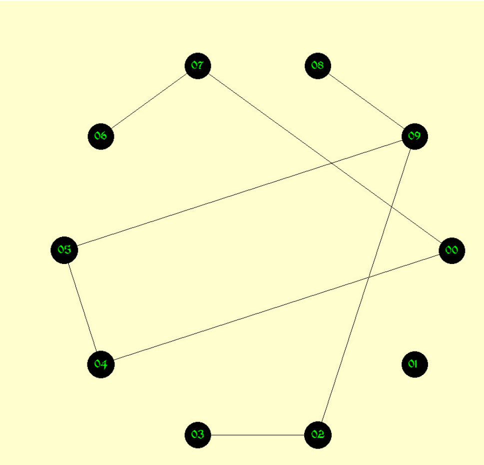
Introduction

- One model introduced by Paul Erdös and Alfréd Rényai in 1959.
- A similar model was introduced at the same time by Edgar Gilbert.
- $G(n, e)$ is a graph chosen at random from all possible graphs with n nodes and e edges.
- $G(n, p)$ is the graph with n nodes and edges choses from all possible edges with probability p.

The Algorithm

- The algorithm I used to implement the Erdös - Rényai model is:
    1. Generate n nodes, where n is determined by the user.
    2. Generate a set of edges corresponding to all the possible edges in K_n.
    3. For each edge, generate a uniform random number $[0, 1)$ and if this number is less than the probability p, also supplied by the user, add the edge to the graph.

An example of $G(10, 0.15)$



Properties of the graphs produced

- On average $G(n,p)$, all nodes have degree close to $(n-1)*p$
- the value $\frac{lnn}{n}$ is a sharp threshold for connectedness. If p is below this value then there will be disconnected components, while above this value the graph will be connected. [W2016a]
- For a 10 node graph, this value is approximately 0.23

Outline

Introduction

- A problem with the Erdös - Rényai model is that all nodes have roughly the same degree.
- Many systems (both natural and man-made), for example social networks, Internet connectivity and citation networks do not follow this pattern.
- Barabási and Albert proposed a model in 1999, with their work on degree distribution on the web.
- Their model is an example of "preferential attachment"
- The idea of "preferential attachment" appears to date back to at least 1925.

The Algorithm

- The algorithm used is:
  1. Begin with a connected network of m_0 nodes. I do this by generating a spanning tree on K_{m_0}.
  2. Proceed to add new node, v_j to the network one at a time and make $m \leq m_0$ connections to the nodes in the original graph.
  3. The probability of an edge being formed between the new node, v_j and an existing node, v_i is given by $p_i = \frac{d(v_i)}{\sum_{v_j \in V(G)} d(v_j)}$
  4. Repeat steps (2) and (3) till all the nodes have been added to the graph.

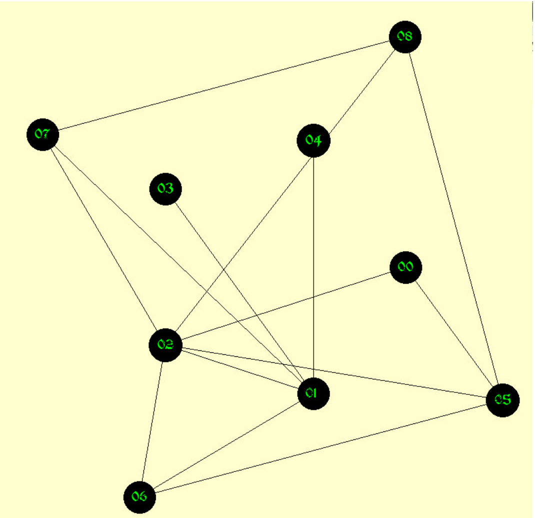
The Algorithm

- A na ive implementation of the above algorithm lead to some problems.
- The constraint on the number of edges to add is ill-defined. In other words, how to decide how many edges to add? I took the expedient way out, I hard coded the number of edges at 3.
- Just generating a uniform random number in [0,1) and comparing it to the probability in step three on the previous slide did admit isolated nodes.
- On further examination (and research), it appears that the best way to do this is to use a normalize degree distribution. [W2015]

Properties of the graphs produced

- The degree distribution is 'scale-free'.
- A 'scale-free' network is a network whose degree distribution follows a power law. In other words, the fraction, $P(d)$ of nodes having d connections to other nodes is given by $P(d) \alpha d^{-\gamma}$, where $\gamma \in (2, 3)$

Example of m_0=5, n=5



Outline

Introduction

- Erdös - Rényi graphs tend to have short path lengths and low clustering.
- Certain 'small-world' networks were observed to have both short path lengths and high clustering [W2016b]
- Small-world networks are networks where nodes are adjacent to only a few nodes, but most nodes are reachable from any other node by a few steps.
- Small-world networks have the typical distance between arbitrarily chosen nodes to be proportional to the number of nodes in the network
- The Watts - Strogatz model is designed to produce graphs that exhibit 'small-world' properties. [wikipedia], *i.e.* $d(u, v) \alpha log(n)$

The Algorithm

- The Watts - Strogatz algorithm is:
  1. Given a number of nodes *n*, an average degree, K (assumed to be even), and a 'special parameter' $/beta$
  2. Generate a circular lattice of *n* nodes, where *n* is the degree of the graph to generate.
  3. Attach each node to the K/2 neighbors going clockwise and counter-clockwise
  4. For every node n_i, take every edge $(n_i, n_j) i < j$ and rewire with a probability $/beta$. Rewiring is done by replacing (n_i, n_j) with (n_i, n_k) where k is chosen at random from all nodes that avoid multiple edges and self-loops.

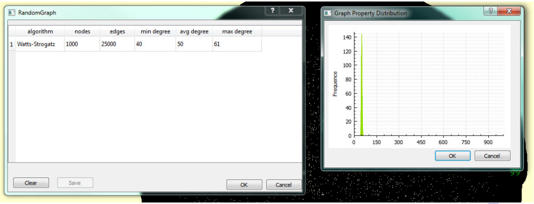
Properties of the graphs produced

- The parameters are expected to follow $n \gg k \gg ln(n) \gg 1$

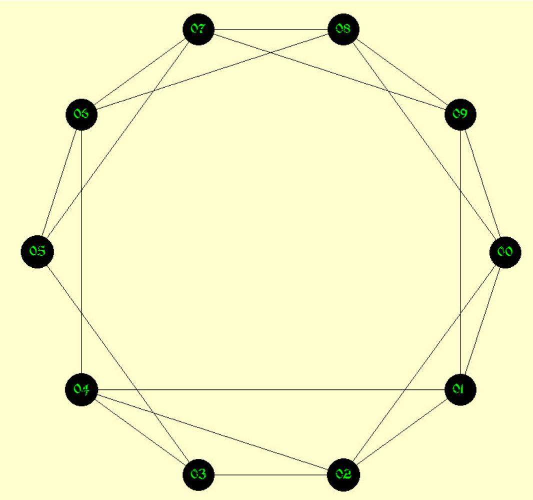
W-S Watts and Strogatz found that L is approximately $\frac{n}{2k} \gg 1$ and the clustering coefficient approaches $\frac{3}{4}$ as β approaches 0.

W-S Watts and Strogatz showed that L approaches $\frac{ln(n)}{ln(k)}$ and C approaches $\frac{k}{n}$ as β approaches 1.

- We expect $\beta \frac{nK}{2}$ non-lattice edges



Example graph with n=10, k=4, b=0.2



RandomGraph application

About the application

"RandomGraph" is an application that allow for visualization of random graphs generated by the three models discussed herein.

- Coded in C^{++} using Qt (https://www.qt.io/), used version 5.1.1
- Consists of 27 files, 14 classes and ~4500 SLOC
- Code compiles with Visual Studio 2010, Service Pack 1.
- Code compiles with gcc, version 4.8.4 on Ubuntu 14.04 LTS.

RandomGraph application

About the application

External Dependencies (both are included in the source tree:

1. *pugixml* Used for reading XML files (http://pugixml.org/), used version 1.7.
2. *QCustomPlot* Used for generation of histograms (http://qcustomplot.com/), used version 1.3.2.

Where the code lives

- Code will be uploaded to a git-hub repository by weeks end.
- Source for the slides will be included as well (used $\LaTeX$, with the *beamer* package to produce slides.)
- I will send out mail with the URL

Summary

- I introduced the concept of random graphs.
- I discussed three models of random graphs, and their potential uses.
- I demonstrated the application created to visualize and explore random graphs.

References

- Erdös, P; Rényi,A. "On Random Graphs 1". *Publicationes mathematicae* **6**: 290-297 (1959)
- Gilbert, E.N. "Random Graphs". *Annals of mathematics Statistics"* **30**(4):1141-1144 (1959)
- Watts, Duncan J.; Strogatz,S.H. "Collective dynamics of 'small-world networks' *Nature* **393**(6684): 440-442
- Barabási, Albert-Làszlò. Network Science. Cambridge University Press:Cambridge. 2016 (Chapter 5 available online at http://barabasi.com/f/622.pdf, last acessed 17Jul2016)
- Chung, Fen. "A whirlwind tour of random graphs" (available at: http://www.math.ucsd.edu/ fan/wp/randomg.pdf, last accessed 17Jul2016)

References

W2016a https://en.wikipedia.org/siki/Erdös-Rényi_model (last accessed 17Jul2016)

W2016b https://en.wikipedia.org/wiki/Watts_and_Strogatz_model (last accessed 17Jul2016)

W2016c https://en.wikipedia.org/wik/Barabási-Albert_model (last accessed 17Jul2016)

W2015 https://compuzzle.wordpress.com/2015/02/03/generating-barabasi-albert-model-graphs-in-clojure/