M1 Writeup

March 4, 2015

Contents

1 Structure generation

1.1 Covariates

Define X to be a $n \times p$ centered and scaled covariate matrix. The covariates are multivariate normal $MVN_p(\mu, \Sigma)$ where $\mu_i = 2\widetilde{C}_{i,1}^2 + 1\widetilde{C}_{i,2}$ and $\widetilde{C}$ are the centered and scaled coordinates.

The covariance matrix has the structure $\Sigma_{ip+s,jp+t} = \rho e^{(-\tau E_{i,j} - \eta|s-t|)}$. Thus, the variance of the individual covariates are 1 and the correlation decays as $E_{i,j}$ increases and as $|s-t|$ increases.

We need the mass of each individual node to compute the gravity term in the gravity model. In the case of White-Nose Syndrome, the mass is the number of caves. Arbitrarily set the first covariate to be the mass of the node. Define $M_i = \lfloor X_{i,1} - X_{[1],1} \rfloor + 1$ to be the mass of location i.

1.2 Network distance

In order for the structure of the network to have a strong impact on the disease spread, we define distance as the geodesic distance raised to a power. The distance between adjacent nodes (i and j such that $A_{i,j} = 1$) is defined to be 1. For non-adjacent nodes, the distance is the number of edges along the shortest path raised to a power. To determine the appropriate power, let $p_{i,j}^{(1)}$ be the hypothetical probability node i infects node j where i and j have geodesic distance 1. Let $p_{i,j}^{(2)}$ be the hypothetical probability if infection if i and j have geodesic distance 2. For their mass, use the mean mass of the network. The log odds-ratio is equal to

$$-\frac{\alpha}{m^\rho} + \frac{\alpha 2^z}{m^\rho}.$$

Set this equation equal to $\log(0.5)$ and solve for z.

1.3 Structures

1.3.1 Random Network

The random network generates location coordinates uniformly on $[0,1] \times [0,1]$. Let C denote the set of coordinate pairs. For any $i \in \{1, ..., n\}$, $C_i \in [0,1] \times [0,1]$.

Let A be the $n \times n$ adjacency matrix where $A_{i,j} = 1$ if locations i and j are adjacent and $A_{i,j} = 0$ otherwise. To determine which locations are connected, we must first define some notation. Let $E_{i,j} = \|C_i - C_j\|_2$ and $E_{i,[j]}$ denote the j^{th} largest norm excluding $E_{i,i}$ which is zero. The set of k nearby locations to location i is $N_i^{(k)} = \{j : E_{i,j} <= E_{i,[k]}\}$. Now, the adjacency matrix is defined as $A_{i,j} = 1$ if $i \in N_j^{(k)}$ or $j \in N_i^{(k)}$.

1.3.2 Starred Alley

The starred alleyway was produced to incorporate the importance of network centrality metrics. An example of a starred alleyway can be found in figure 2. By looking at the structure, the goal

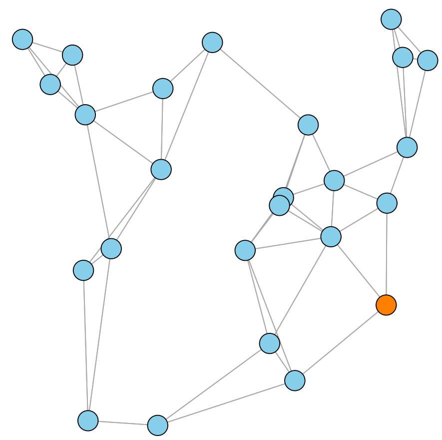

Figure 1: A random network with 25 locations

is to identify the locations along the main horizontal line. If these locations are treated, it helps to contain the spread of the disease.

Distances between nodes are not defined in the usual way. For any node on the outer points, the distance to the closest node on the main horizontal line is 0.9 instead of 1.0.

For a starred alley of size N, define

$$m_N = \arg\max_{m>0} f(m)\mathbb{1}_{\{f(m)\leq N\}}$$

where

$$f(m) = m + \left\lceil \frac{m}{2} \right\rceil \left(\left\lceil \frac{m}{2} \right\rceil - (m \bmod 2) + 1 \right).$$

Let locations $1, \ldots, m_N$ index the locations on the main horizontal line. Next define disjoint sets of locations $\mathcal{N}_i$ for $i = 1, \ldots, m_N$. The cardinality for $\mathcal{N}_i$ is defined as

$$|\mathcal{N}_i| = \left\lfloor \frac{i}{2} \right\rfloor + \mathbb{1}_{\{(m_N - i) < (N - f(m_N))\}}.$$

For each location $j \in \mathcal{N}_i$ set $A_{i,j} = A_{j,i} = 1$.

1.3.3 Ring

The ring is a simple circle of locations with a concetrated group in one section. To form the ring, place the locations evenly on the circle then between one pair of locations insert more locations

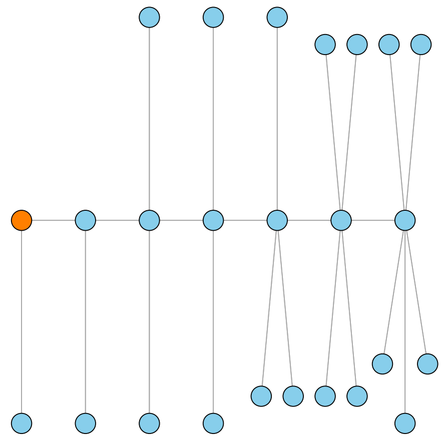

Figure 2: The starred alleyway network with 25 locations

evenly. An example can be seen in figure 3.

1.3.4 Grid

The grid is a checkerboard-like structure.

1.3.5 Bowtie

The bowtie combines two grid networks with a random network in between. Spread of the disease begins in one grid and travels through the random network and into the other grid.

1.3.6 Scalefree

The scalefree network is designed so that some locations are highly connected while others are only connected by a couple edges. This type of network was originally desigined to mimic the network of webpages. The internet houses many websites that are highly connected (e.g. Google, Yahoo, etc.) and others that are not.

To generate the network, we will define the process via induction. A network of size 1 is the trivial network. Given a network of size $N-1$, attach the additonal N^{th} node to existing node

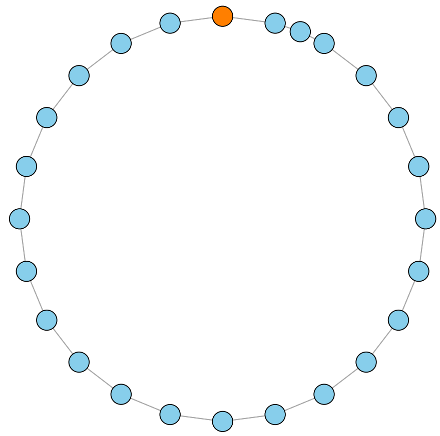

Figure 3: The ring network with 25 locations

$i \in \{1, \ldots, N-1\}$ with probability

$$\frac{\sum_{j!=i} A_{j,i}}{\sum_{j,k\,:\,j>k} A_{j,k}}.$$

Generating a scale-free network in this fashion results in a Barabasi-Network. The defining feature of a scale-free network is the proportion of nodes with k edges is proportional to $k^{-\gamma}$ for some γ. In the case of a Barabasi-Network $\gamma = 3$.

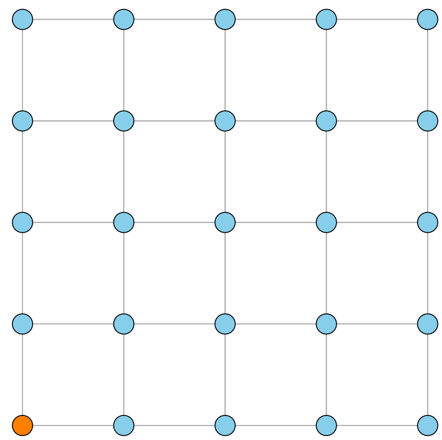

Figure 4: The grid network with 25 locations

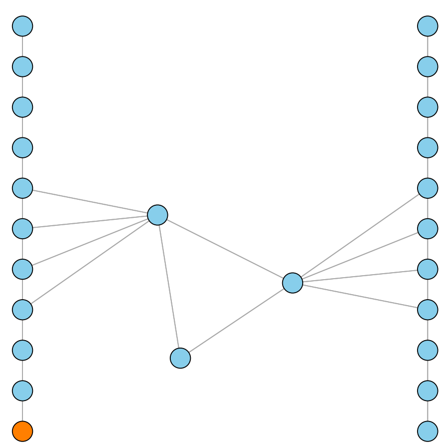

Figure 5: The bowtie network with 25 locations

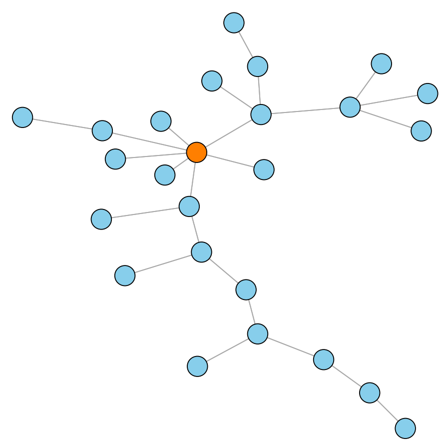

Figure 6: The scalefree network with 25 locations