Research Proposal
Jason Raiti
Dr. Richard Bertram

Agent-based analysis of functional hubs

INTRODUCTION

The modeling of networks is a field of math with many applications across science. One of these areas is in that of biological neural networks. Another of these areas is in disease modeling. Although practically different, these networks have many functional similarities, and understanding of one can aid in understanding of the other. A question that is relevant to most network applications is how a signal is transmitted through the network, and what properties of specific nodes in the network make them crucial in the transmission of the signal. That is, what makes them "functional hubs" of the network. In respect to the effect of the physical arrangement of the network, the existence of hub nodes (with disproportionate connectivity to other nodes) encourages the spread of activity or disease between nodes. These are called "structural hubs". Are all structural hubs functional hubs? Are all functional hubs structural hubs? These are questions that we will investigate using an agent-based model.

RESEARCH METHODOLOGY

Begin by modeling SIR models with different network dynamics. This begins with programming with Erdós-Rényi (random) adjacency matrices, and modeling an SIRS model within iterations of these networks. This will be a training exercise for building network models and implementing agent-based models. Research on hubs themselves will be more general than epidemic models.

The next step will be to stimulate a more general threshold process on an E-R network. Here, the state of each node is 0 unless then sum of neighboring state values is greater than some threshold θ. If it is, then the respective node's state is set to 1. This model will be a way to simulate beta cells in a spiking state (1), and not in a spiking state (0).

The final step is to construct networks such that there will be functional hubs in the implementation of the threshold process. This may come in the form of structural hubs or in the form of bridges. Bridges are defined as connections between hubs that are part of completely independent or distantly connected "cliques" of hubs. Visualization will be aided using the Gephi software package. This part of the research seeks to investigate what are the properties of these functional hubs and must they also be structural hubs?

RESEARCH QUESTION

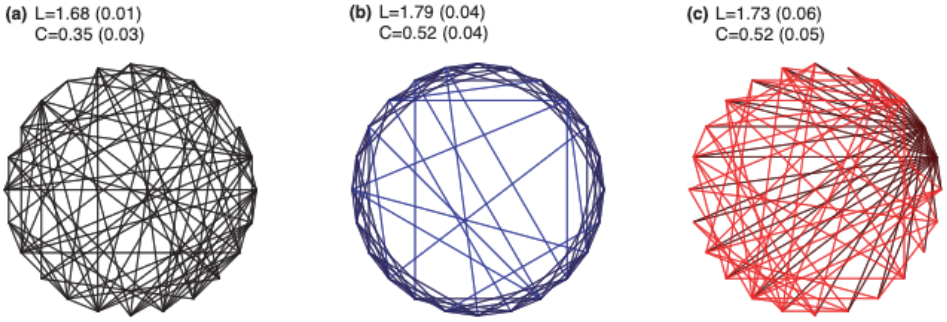
What properties of a node in a network make it a functional hub? Are functional hubs also structural hubs or can they be non-structural hubs?

BACKGROUND AND RATIONALE

This research will be extensively using agent-based models to investigate the role of network architecture on the dynamics of interactions within a population. With these models, the state of each node is updated at each time step according to its current state and the state of neighbors (where a neighbor is a node connected to it by an edge). This state update can be due to a probability constant (SIR model) or by a threshold constant (threshold model).

The goal of using an agent-based model is to gain insight into the collective behavior of individuals (nodes) obeying simple rules like in natural systems. This is in contrast to traditional compartmental models, which seek to describe the system as a whole through equations, which generally describe the sum of all interactions in a system. Agent-based models can be thought of as micro scale models that attempt to recreate the process of emergent behavior in a system. Where compartmental models attempt to describe the macro-scale behavior of a system, agent based models attempt to show how the macro-scale behavior emerges from simple micro scale interactions between individuals to create complex whole system level state changes. Thus, agent based models have the advantage of accounting for nuances in systems which compartmental models may overlook.

Agent based models are based on networks. Networks are sets of nodes linked by connections or edges. These can be mathematically described as graphs. These nodes may represent individuals of any type: neurons, people, web pages, etc. Generally networks can be described as random (Erdós-Rényi), small world, or scale free.



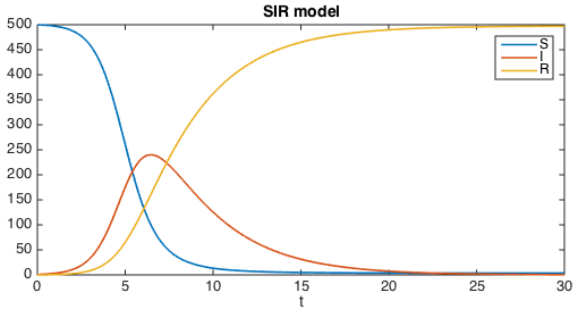
The above diagram shows examples of different network architectures. L represents average path length between nodes. C represents the clustering coefficient. Clustering coefficient is the average number of nodes any gives node is connected to divided by the number of possible connections. The figure shows the structure of random, small-world, and scale-free networks respectively from left to right. All networks in the figure have 24 nodes and 86 connections with nodes arranged on a circle. (Sporns, 2004)

The classic example of an agent-based model is that of the SIR model. This is the reason this research will begin with building SIR models, and will adapt the SIR model to simulate free-scale models of pancreatic beta cells.

An SIR (Susceptible-Infected-Recovered) model is an epidemiological model that represents the number of people infected in a closed population over time. It is meant to be a rough approximation of the lifetime of a disease in individuals that can have brief immunity after infection.

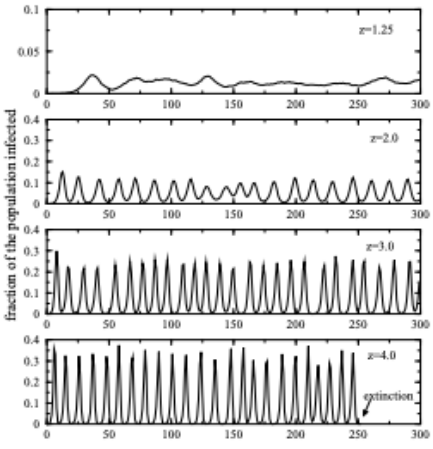
SIR can be implemented as a compartmental model as well, where members of the population are described by their average rather than accounting for their individual interactions. The system dynamics of this compartmental SIR model are then described by ordinary differential equations.

$$\frac{dS}{dt} = -\frac{\beta IS}{N},$$
$$\frac{dI}{dt} = \frac{\beta IS}{N} - \gamma I,$$
$$\frac{dR}{dt} = \gamma I,$$



In this system of equations, the fraction of the population susceptible to a disease, infected by disease, and recovered with immunity is denoted by S, I, and R respectively. β is the average number of contacts per person per time, multiplied by the probability of disease transmission in a contact between a susceptible and an infectious subject. This constant represents the generalized probability of infection between any two individuals in the infected and susceptible states. γ is the transition rate between the state of being infected and being recovered.

An interesting feature of SIR models is that if another constant is added to account for transition from the recovered state to the susceptible state, emergent behavior can be observed at the macro scale. This modified SIR model (called SIRS) can display regular synchronized oscillations in the number of individuals infected. The figure to the right is an example of this synchronized oscillation. (Girvan 2002)



These models are limited in that they assume perfect mixing of a population when in reality individuals have a finite number of connections with other individuals in a population.

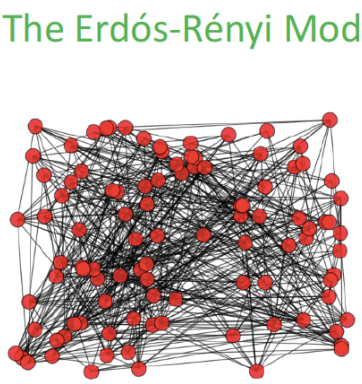
To implement this model as an agent-based model, the nodes in the network would all begin in the susceptible state. Then, a set number of randomly chosen nodes will be “infected”. For each time step in the model, infected individuals will have a constant percent chance of recovering, and susceptible individuals with connections to nodes that are infected will have a constant percent chance of being infected.

The agent-based model is a much better way to describe the spread of disease through a population, as models that do not account for the spatial interactions of individuals have been shown to be less accurate (Eisinger 2008). At each point in time, the state of a node is based on its own state and that of its neighbors. This implementation of the SIR model will be the starting point for the implementation of a more general threshold model on a network.

This agent-based model will be implemented on a random network, also known as an Erdós-Rényi network. In such a network, each of the n number of nodes connect with each other node with a fixed probability p. This graph is thus denoted as G(n,p). Erdós-Rényi networks typically do not allow for self edges. The advantage of these graphs that they may be generated in many iterations quickly and easily. This type of network can be computationally modeled using an adjacency matrix implemented in scilab.

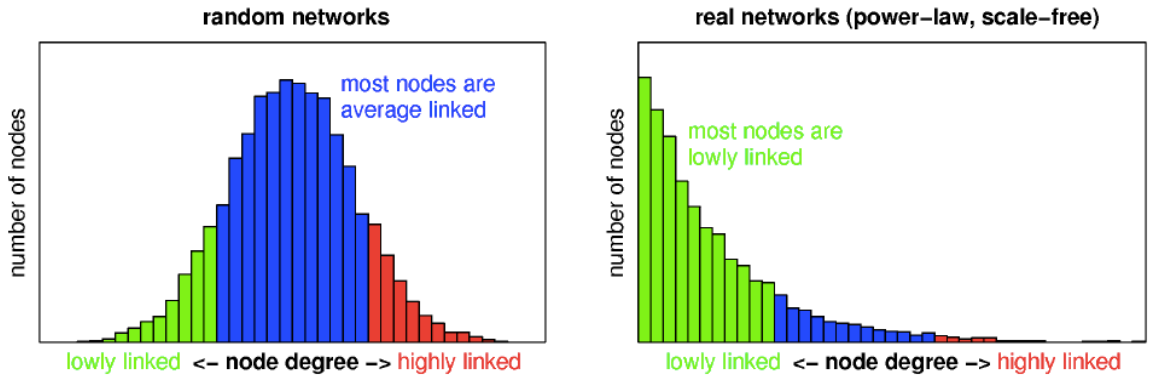
```
"Adjacency Matrix:"

 0.   0.   0.   1.   0.   0.   0.   0.   1.   0.
 0.   0.   0.   1.   1.   0.   0.   0.   1.   0.
 0.   0.   0.   0.   0.   1.   0.   1.   1.   0.
 1.   1.   0.   0.   0.   1.   1.   0.   0.   0.
 0.   1.   0.   0.   0.   0.   0.   0.   0.   0.
 0.   0.   1.   1.   0.   0.   0.   0.   0.   0.
 0.   0.   0.   1.   0.   0.   0.   0.   0.   1.
 0.   0.   1.   0.   0.   0.   0.   0.   1.   1.
 1.   1.   1.   0.   0.   0.   0.   1.   0.   0.
 0.   0.   0.   0.   0.   0.   1.   1.   0.   0.
```



The figure above is a ER adjacency matrix and a example of a graphical representation of an ER network. In the matrix above, zeros represent no connection, and ones represent connection between nodes. The ER network can be generated by first building an empty zero matrix of size n by n. Using the rand() function compared to threshold p, the connections can be randomly assigned. This process can be repeated to generate many unique ER networks all of the same G(n,p). Thus, it is possible it iterate unique networks to run the SIR simulation on.

The behavior of SIR modeled epidemics can be compared with properties of ER models like degree distribution, mean degree, and mean number of edges. This is possible because for ER graphs, a property like degree distribution is fixed, and all other features (like connectivity pattern) are determined randomly. This makes it possible to characterize an entire population of networks rather than a single network.

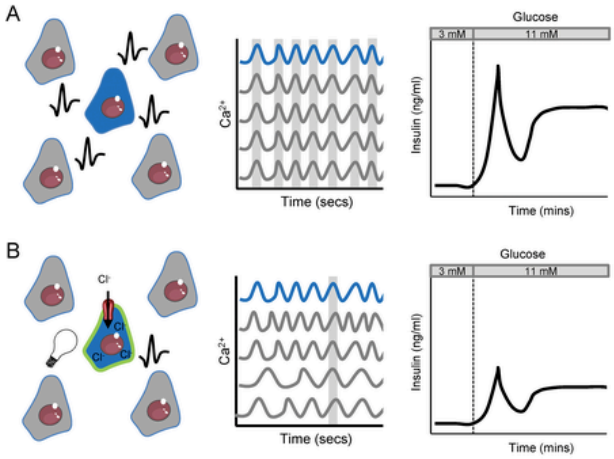


It is important to note that there are limitations how representative ER networks are. Random networks tend to have a normal distribution in node degree; while in reality most networks have a power law node degree distribution. The ER network can be manually modified be more representative of small world and scale free networks. These networks are also of interest because of they display enhanced signal-propagation speed, computational power, and synchronizability. (Watts 1998)

In addition, the SIR model is not an ideal model for analyzing the role of hubs due to the randomness in transmission. A threshold model is an alternate method of investigation and can be implemented on ER networks.

The most famous example of a threshold network is Conway's Game of Life. In such a network the state of each node is 0 unless the sum of connected node state values is greater than some threshold θ. If the threshold is passed then the respective node's state is set to 1. This removes the random transmission constant of the SIR model, making analysis of the effects of network architecture clearer. In addition, the threshold network is a better simulation of the biological networks this research seeks to better understand. A value of 1 can be thought of as a neuron or beta cell in a spiking state, a value of 0 means non-spiking state. The threshold network also can develop synchronization.

In respect to the behavior of real world pancreatic beta cells, there are highly interconnected hubs, which act as pacemakers that are connected to many follower cells. These hub cells compromise of ~1% - 10% of the total beta cells. This architecture is important in efficient insulin secretion. Loss of connectivity due to gene deletion is associated with type 2 diabetes.



The figure is a demonstration of this behavior. “A” shows the normally functioning network, with grey follower cells and blue hub cells. This leads to coordinated synchronized behavior and normal insulin release. “B” shows a hub cell that has been optically silenced using halorhodopsin. Network behavior is more stochastic due to this altered hub behavior and as a result insulin secretion is impaired. (Rutter 2017)

Thus there is the possibility of restoring normal function to certain cases of diabetes through a better understanding of how to restore synchronization in the pancreatic cell networks.

BIBLIOGRAPHY

Eisinger, D., & Thulke, H. H. (2008). Spatial pattern formation facilitates eradication of infectious diseases. The Journal of applied ecology, 45(2), 415–423. https://doi.org/10.1111/j.1365-2664.2007.01439.x

Girvan, M., Callaway, D. S., Newman, M. E., & Strogatz, S. H. (2002). Simple model of epidemics with pathogen mutation. *Physical Review E*, *65*(3). https://doi.org/10.1103/physreve.65.031915

Rutter, GA, Hodson, DJ, Chabosseau, P, Haythorne, E, Pullen, TJ, Leclerc, I. Local and regional control of calcium dynamics in the pancreatic islet. *Diabetes Obes Metab*. 2017;19:30–41. https://doi.org/10.1111/dom.12990

Sporns, O, et al. "Organization, Development and Function of Complex Brain Networks." Trends in Cognitive Sciences, vol. 8, no. 9, 2004, pp. 418–425., doi:10.1016/j.tics.2004.07.008.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, *393*(6684), 440-442. https://doi.org/10.1038/30918