Network Characteristics’ Effect on Cascade Propagation

In investigating the spread of cascades, and the effectiveness of halting cascades via partial vaccination with limited knowledge of the network, it may be helpful to understand factors that affect cascade spread, both positively and negatively. A cascade is potentially a network-wide event, as can be seen by the prevalence of ‘fake news’ spreading through social networks, and malware spreading to millions of smartphones or computers. However, a cascade is also a node-level event. An individual node becomes infected not through large-scale network-level factors, but, for the most part, through interactions with its neighbors. Therefore, it may prove useful to examine both large- and small-scale characteristics of a network, and their effects on cascade spread.

Node Characteristics and Effects on Cascade Spread

Since nodes are directly infected by their neighbors, and in turn pass on the infection to other neighbors, a brief overview of node characteristics and cascade spread is in order. The most obvious node characteristic to examine is degree. A high-degree node has more chances to infect other nodes and, depending on the details of the infection model used, has more chances to become infected. However, depending on location in the network, this may or may not always be pertinent. A high-degree node on the periphery of the network may not have much effect on cascade propagation.

Considering that a cascade is the flow of information through a network, perhaps a more useful characteristic to focus on is betweenness centrality. Higher betweenness indicates that a node represents a common link in multiple transmission paths, and the ability to discover and vaccinate such a node could be effective in halting cascade propagation.

Since ours is an undirected, weighted graph, with weights representing duration of close contact, the cascade spread, and prevention, are more complicated than a singular measurement can reflect. A node with few heavily weighted edges may be a more prolific infector than is a node with very high degree but lightly weighted edges. Choosing a vaccination strategy with these considerations in mind may prove to be a difficult problem.

Network Characteristics and Effects on Cascade Spread

Intuitively, one can see that a complete graph G=(V,E), with *n* nodes and *n(n-1)* edges provides an environment very conducive to cascade propagation. As a simplistic example, consider a crowded classroom. With one sneeze by an infectious individual, the entire population is potentially infected. On the other hand, a completely disconnected network with *n* nodes and 0 edges makes it impossible for the cascade to spread to even one node beyond the seed.

Obviously, real-world networks lie somewhere in between these two extremes. Historically, research into epidemic spread assumed a homogeneous contact model of nodes in a population, making it possible for any node to infect any other node, similar to the complete network example above. However, the presence of disconnected components, small-world characteristics, clustering, and modularity, among other factors, can act to encourage or inhibit cascade propagation.

Since the focus of this project is a real-world contact network in an American high school, some of these factors are more pertinent than others. Disconnected components are the strongest barrier to propagation spread, but possibly the most difficult to exploit in our network. Whereas an elementary school, for example, may have many disconnected components due to class and grade separation, a typical American high school has a very high degree of mixing due to change in classroom population from class to class, periods between classes, and club or team memberships.

The small-world effect is another concern to our project. Our target network does exhibit small-world properties [1], which will aid in cascade propagation. We expect that, depending on probability of infection, the cascade will spread to a large percentage of nodes quickly if effective immunization is not implemented. Modularity may act to increase the length of the cascade if immunization is improperly implemented, but may also aid in containment, depending on the modularity of the neighborhood to which the seed node belongs.

Some network types are naturally more resistant to cascade propagation than others. An example of such networks is the ‘dark’ network, such as a crime ring or terrorist group. Such networks are designed to limit spread of a cascade if an individual node becomes infected (detained by law enforcement, for example). Networks with low average degree and high modularity have been determined to be the most resistant to cascades [3]. Disconnected networks, such as a disconnected star graph or disconnected cliques, make the likelihood of a network-wide cascade impossible. Of course, this type of network is not a feasible model for design of most real-world networks. However, it may be possible to utilize targeted vaccination to induce higher modularity in a network.

From an information cascade perspective, an immunized node is effectively removed from the network, since travel along an edge to an immunized network is impossible. With targeted immunization, it may be possible to effectively create disconnected networks with zero probability of wide-spread infection, from an otherwise cascade-vulnerable graph.

Modularity Findings:

RandomVac:

Since randomized vaccination forms the baseline by which other partial knowledge, partial vaccination strategies will be judged against, it is useful to examine the effect of random vaccination on modularity from a cascade perspective. Since vaccinated nodes are effectively removed from the network as far as a cascade is concerned, it is possible to track the changes in effective modularity of the network as vaccination progresses. Random vaccination, unsurprisingly, produces no noticeable trend in modularity change until nearly all nodes are immunized (removed from possible cascade paths). As the cascade-susceptible network shrinks to a small fraction of the original network, the modularity score rises erratically.

modularityStuff/randomVacModularity/randomVacModularity.pdf

RandomWalk:

Effective modularity score of the network when a random walk vaccination strategy is applied with partial network knowledge shows similar behavior to the random vaccination strategy. However, the increase in effective modularity occurs at a smaller fraction of vaccinated nodes than in the random vaccination strategy, and the peak effective modularity score is higher. In a random walk vaccination strategy, a node is chosen at random for vaccination, then a neighboring node is chosen for vaccination. This walk continues until the desired number of nodes has been vaccinated. The major difference between this vaccination strategy and random vaccination is that, in the random walk, vaccinated nodes are connected. In a large network with a short vaccination walk, large changes in effective modularity do not occur. However, as the number of immunized nodes increases, this immunization walk may act to cut inter-household paths, thereby increasing modularity. Unfortunately, this effect isn’t seen in our network until the number of vaccinated nodes passes 600. This is sooner than that seen in the random vaccination strategy, but is still nearing the point of full vaccination.

modularityStuff/randomWalkModularity/randomWalkModularity.pdf

MaxStrength:

In contrast to the maximum degree score, the maximum strength score measures the sum of edge weights connecting to a node, rather than number of edges. This measurement is pertinent to our study because of the weighted edges in the network, and particularly because transmission probability between nodes is proportional to the weight of the edge linking the two nodes. A node with a few heavily-weighted edges is more likely to spread the infection than is a node with lightly weighted edges. Of course, a node with an extremely high number of lightly weighted edges could have the highest maximum strength score. However, the degree distribution of our network appears to make cases such as these outliers, rather than the norm.

Examining the results of the maximum strength vaccination strategy with respect to effective modularity, a much more pronounced trend is visible than in the random vaccination and random walk strategies. A definite positive trend in effective modularity appears at approximately 450 nodes vaccinated. While still representing a large fraction of the network, this represents a smaller proportion of nodes than was necessary in random vaccination or random walk strategies, and the trend in effective modularity was much clearer in the maximum strength strategy. The smoothness of the trend line reflects the deterministic nature of maximum strength vaccination, as opposed to the random nature of random walk and random vaccination.

modularityStuff/maxStrengthModularity/maxStrengthModularity.pdf

Betweenness Centrality:

A vaccination strategy targeting nodes with high betweenness centrality scores in a partial knowledge environment yields similar results to a maximum strength strategy, with the main difference being peak modularity in the betweenness strategy is higher than the maximum strength strategy. Again, increase in modularity does not occur until the number of vaccinated nodes approaches 500. This was surprising, considering the relationship between nodes with a high betweenness score and the small-world property exhibited by the network. By vaccinating nodes with high betweenness centrality, we had expected to be able to induce high modularity, and high cascade resistance, at a smaller fraction of node vaccination.

modularityStuff/centralityModularity/CentralityModularity.pdf

[3] A. Gutfraind, "Optimizing Topological cascade resilience based on the structure of terrorist networks," *PLoS ONE*, vol. 5, no. 11, p. e13448, Nov. 2010.