Telegraphic Presentation: Article Question Answers

Gabe, Oj, Jacob

a) The authors modelled disease transmission through a partially connected network of people. This social network was represented by a graph with contact edges between people likely to infect one another. Using this representation, the author hoped to model the effects of fine-scale network structure, in particular of local “cliques” - small, highly connected clusters - on disease spread. The author finally queried the model about various disease-quantities - the threshold for epidemic, the size of the final infected population, and the minimum epidemic-preventing vaccination strategy.

b) Since this particular model was concerned with disease spread on a graph, the author first derived several useful graph-counting expressions for nodes, pairs, and triples of any given type (S/I/R). In order to focus attention on local structure, the author defined the fundamental quantity as the ratio of connected triangles in the graph to triples of any type (including linear arrangements). For a given number of nodes and connections, a higher represents a tendency for people to “clump” into highly connected groups (or triangles), with long range connections between groups being less common. The author also derived approximate formulas for counting triples in the graph in terms of singles, pairs, and the ratio . These were later used to “close” the system in terms of pairs.

With these expressions in hand, the author first wrote down formulas for the derivatives of different pair counts ([SI], [SS], [SR], etc.) in terms of triples, pairs, a contact parameter , and a recovery rate g. This was done simply by counting contacts and nodes, and then assuming that each contact between a susceptible and infected node caused infection with rate and that a fraction, g, of the infected nodes recovered per unit time - the same assumptions made by the simple mean-field model, with the only difference being in the number of contacts between each population, which in this case was a function of the graph structure. At this point, the author ran several stochastic simulations with a large graph (N=6000) and found that the simulated sample paths were, on average, in good quantitative agreement with the model equations.

Later on, to facilitate analysis, the author substituted his expressions for triples into the evolution equations, leaving the model in terms of network pair counts, disease parameters, and the ratio .

In constructing and analyzing the model, the author used several approximations, but there were fewer model-restricting assumptions. One assumption the author did make, however, was in approximating the number of triples in terms of numbers of pairs and singles. Specifically, it was assumed that for each node, the neighboring node states were distributed multinomially - which is equivalent to assuming that given a particular node and its state, each neighbor’s state is independent of all other neighbors’ states. On average, this might lead to expressions which are close to correct. But one can imagine situations, especially early on when an epidemic is beginning to spread, when neighbor states are very non-independent: if, for example, the epidemic starts on the left side of the graph and you’re a susceptible node, knowing your first neighbor is infected makes you more likely to be on the left side, which in turn makes you more likely to have other infected neighbors, which suddenly makes your neighbors non-independent.

c) The author did run some simulations, and these may have provided the hint which caused him to examine node correlations in later analysis. However, most of the results were obtained by analysis of the model. For example, in order to get at the threshold for epidemic and the final size of the infected population, the author mainly studied the limiting behavior of the model equations. He also analyze the effect of making the initial susceptible population smaller - “vaccinating” - on disease spread, and derived an approximate vaccination threshold for prevention of epidemics. In all cases, simulations were used primarily for visualization rather than for quantitative insights.

d) The author’s main message was actually somewhat intuitive: for a fixed average number of neighbors, a higher connectedness generally made the final infected population lower, epidemic threshold higher, and vaccination threshold lower. This is probably because if the average number of neighbors is held fixed, then as is increased, the graph becomes much more “clumped”, with fewer and fewer connections between clumps. While it is easy for diseases to spread through each connected clump, the limited contact between clumps acts as a sort of bottleneck that slows spread. In effect, with increasing , each clump is increasingly sequestered. This is an interesting and intuitive message: populations with local “tribes” which only connect to the rest of the population in a limited way, are more resistant to disease.

e)

i) Though I had a hard time reading the article, I can’t say it wasn’t well written, and considering it’s intended audience, clear enough. Steps were skipped in the analysis, but for readers familiar with the literature (not us), the missing links may have been obvious. The author’s approach was also methodical and logical: first he derived graph counting expressions which would be useful later; then he wrote down a model; then he used his counting expressions to simplify the model; and finally he analyzed the model.

ii) The only assumption that I was able to find - about neighbor states being distributed multinomially - seemed as though it could’ve used more justification. It also would have been useful to know how much of the model’s behavior depends on this assumption, since this is hard to guess a priori.

iii) Since the author was interested in how incomplete connectedness affects disease spread, I’d say this is a very good model, which includes the relevant details - simple infection and recovery dynamics and graph structure - and not much more. At the end, we are left with some very satisfying relationships which seem to capture the effects of the connectivity parameter in a way that is consistent with our qualitative intuition.

iv) Although it would be difficult to model, it might be interesting to study the effects of variable disease susceptibility among individuals. For example, it’d be fascinating to demonstrate either analytically or through simulations the intuitive fact that even a small fraction of very susceptible people - infectious “weak links” - could greatly attenuate the epidemic protection afforded by clumping, since if these weak links had connections to the outside cliques, they would function as low-resistance entry points into their own highly connected clique.

Presentation Overview:

1. Question: how does a population’s fine-scale contact structure affect the spread of disease?
2. Approach: Author uses a partially-connected graph to represent a population, and studies the spread of disease through the graph as a function of the graph’s “connectedness” .
3. Methods:
   1. Author first derives useful graph-counting expressions and defines graph singles, pairs and triples, connectedness , and node correlation C.
   2. Then, approximate expressions for counting triples in terms of singles, pairs and are derived. These are used later to close the system in terms of pairs
   3. The author then, in a very straightforward way, derives evolution equations using same assumptions of mean-field model - that when an infected and susceptible person meet, the susceptible is infected with some probability (or rate) , and that infecteds recover with rate g. The only difference between the mean field model and this one is in the number of connections between infecteds and susceptibles, which now, thanks to the graph structure, is only a subset of all possible S-I connections. For the evolution equations, the connections are counted in terms of graph pairs and triples.
   4. Author then runs simulations (maybe we can show graphs in paper here)
   5. … then closes system at level of pairs for further analysis
   6. He examines the early evolution of the network initialized with a single infected node. His equations predict that the reproductive ratio of the disease is actually lower than would be expected from the mean-field model (which might correspond to the fact that a disease can quickly spread through the highly-connected “tribe” that it starts out in, but infection slows down substantially when it has to leave the tribe to find more susceptibles).
   7. In the simple case of =0, author finds results similar to (and cites) Kermack & McKendrick. In general, however, author is not able to analytically determine the final size of the population, so he uses numerical integration to obtain results (show figure).
   8. He also incorporates the concept of vaccination in his model by simple converting a fraction of the susceptible population to the recovered state. From this, he derives a vaccination threshold

Take home message: for a given average number of neighbors, author finds lower final epidemic size, lower vaccination threshold, and generally lower population infection when goes up. This is intuitive: as population gets clumpier… blah blah, see above.