

Marissa Graham

CS 670 Midterm 2

I chose to read Jackson's "The Economics of Social Networks", which investigates the motivations behind and advantages of modeling real-world social networks with both random and strategic models. Jackson does not define exactly what he considers to be a social network, choosing instead to give examples, but his working definition seems to be any network whose edges represent the interactions and connections between individual agents.

Random Network Models

In random network models, interaction between agents (network edges) is determined by various stochastic processes. We do not place utility values on agent interactions, and there is therefore no intrinsic notion of behavior dynamics or equilibrium involved. Random models do not explicitly address the question of *why* social networks form as they do, but they do allow us to compare and classify types of networks based on their various structural properties. The benchmark definition for similarity that this provides is a necessary prerequisite for discussing whether the structures resulting from strategic models accurately reflect real-world network structures.

The structural features Jackson notes as intrinsic to social networks are:

1. A high clustering coefficient, that is, on average some of an agent's neighbors are likely to be connected to each other.
2. Low diameter, that is, there is a relatively short path between any two agents.
3. A scale-free degree distribution; that is, the degree distribution fits an exponentially decaying curve.

He introduces three stochastic models for network formation¹, none of which satisfy all three of these properties, and several hybrid models, which interpolate between two of the three formation strategies in order to better match observed networks. These better matches are unlikely, however, to represent a correspondingly better explanation for how the observed networks themselves form; the match is better, but the model is ad hoc and sort of "hard coded" rather than intrinsically matching the desired features

Strategic Models

Strategic models, by contrast, do place utility values on agent interactions, and therefore imply a notion of behavioral dynamics as agents are incentivized to form and sever relationships with other agents. In order to study these dynamics, we modify the notion of a Nash equilibrium to account for the consent of both agents in forming a relationship; as in battle of the sexes, no utility can be derived from non-coordinated actions. Jackson presents three equilibrium concepts:

- Efficiency, essentially "maximize the sum of utility for all agents over all possible connection graphs between them".

¹A standard Erdős-Rényi uniform edge probability model, which has low diameter but fails to be scale-free or clustered, a Barabasi-Albert preferential attachment model, which is scale-free and exhibits very slow diameter growth but fails to cluster, and a Watts-Strogatz ring lattice rewiring model, which demonstrates low diameter and good clustering, but whose degree distribution interpolates between that of an Erdős-Rényi random graph and a Dirac delta function rather than being anywhere near scale-free.

- Pairwise stability, that is, no player wishes to sever an existing link and no two players want to add a link.
- Pareto efficiency, which as usual means that no player’s utility can be increased without decreasing the utility of some other player. Efficiency clearly implies Pareto efficiency, but the converse is not true, and Jackson only discusses the strategic models he presents with respect to efficiency and pairwise stability.

He then discusses the conditions required for and resulting structures of efficient and pairwise stable networks for both a connections model (where links represent social relationships between players, and have benefits and costs, which propagate through indirect links but at a discounted rate), and a business relationships model (where firms forming a link lowers their production costs but indirectly affects the overall market competition structure). In both cases, we can find networks which are efficient and pairwise stable, and efficient but not pairwise stable, depending on model parameters. In the business case, under a Bertrand model of competition between firms (the firms charging the lowest price split the market), there is *no* pairwise stable network which is efficient from either the industry profit or consumer standpoint. As in the prisoner’s dilemma, we find tension between different equilibrium concepts.

In all cases, the equilibrium networks have a highly regular and homogenous structure (complete, empty, star, or interlocking star), which is the result of the strong homogeneity in the payoff functions that allows for a detailed analysis. We are still well-positioned to give insights into the “why” of network formation, as Jackson discusses in Section 3.2.6, but our models are not well suited to predict or match the structure of real-world networks.

Errors and Model Suitability

The main flaw I observed in this paper is that Jackson’s analysis of *why* social networks have certain structural properties is based on the assumption that connections between individuals are generated *independently* of one another, which is not the case for many social networks. It is absolutely not the case for collaboration networks, which Jackson uses as a primary example of a social network. Collaboration networks are one-mode projections of a larger bipartite² network with edges between individuals and the activities in which they participated³—coauthors working on papers, actors working on a film, students taking the same courses, and so on—and the pairwise interactions involved are far from independent. Each activity will result in a complete subnetwork in the larger network, and similar activities (papers on similar topics, courses in the same major, etc.) will correspond to non-complete but even larger clusters of individuals.

It is possible that the combination of high local clustering with short overall path length is still partially explained the equilibrium models described by Jackson (in particular, due to the “islands model” discussed in Section 3.2.6, where connection costs are higher with decreased proximity, but the substantial benefits of a link connecting distant agents results results in a low overall diameter), but it is unlikely that they are the true primary mechanism behind it. He does discuss this somewhat in the conclusion, but only tangentially and offhand, and in the sense of “interesting questions for further study”. I feel that the lack of consideration of independent vs. non-independent link formation weakens Jackson’s claim in Section 3.2.6 that “economic forces tell us a great deal about *why* we should expect to see ‘small-world’ behavior” enough for it to be an error in the paper, rather than a limitation of scope.

²i.e., there are two vertex types, and edges run only between vertices of different type.

³We can represent independent interactions as bipartite networks where each activity has two participants, so the structure of the degree distribution over activities would be an interesting avenue of further study for these models.

When we use strategic models to explain the formation of real-world networks, we must be careful to ensure that we are not making assumptions in our model that fail to accurately reflect the networks we wish to describe, or our conclusions lose their explanatory force. Random models do not introduce this same contingency. We may use heuristics in our model in order to more accurately mimic the structural properties found in real-world graphs, and a simple formation model which accurately reflects many key properties of a real-world network indicates that the formation model itself might also accurately reflect reality, but its usefulness is not dependent on this claim. Models which mimic structural properties also allow us draw connections between networks which are structurally similar despite wildly different origins, and vice versa—why are protein-protein interaction networks, film actor collaborations, and the internet all scale-free, while food webs, coauthorship networks, and power grids are not⁴?

Future Research Directions

A potential next step in this research would be to investigate how well strategic models of interaction explain the structure of a broader variety of networks that exhibit (or lack!) the degree distribution, clustering, and diameter properties which Jackson finds to be characteristic of social networks. In order to do so, we could impose utilities on real-world and random networks according to the strategic models described in the paper, and computationally model the resulting interaction dynamics. If the strategic model is accurate, systems should already be close to a pairwise equilibrium, and will not drastically change their structure. We should be able to gain interesting insights from which networks are already close to a pairwise equilibrium under different models, the types of equilibrium structures we see in networks which are not, and which random and real-world networks have similar behaviors under a certain model.

We should also look further into bipartite graphs and their one-mode projections in comparison to independent-interaction social networks. To what extent does the degree distribution over activities in a bipartite network predict the clustering in the projection? Does it affect the diameter? If we take the projection onto the activities rather than the participants, do we see similar clustering levels and low diameters? What random models specific to the construction of bipartite graphs most accurately mimic the structure of real-world social networks, both overall and in their projections? Can we modify our strategic interaction models to include the entire bipartite interaction graph, and will this give us better results? Future work in these directions could provide many interesting insights, and be a promising method for modeling the heterogeneity in types of interactions and correlation structures that Jackson addresses in the conclusion as inadequately captured by existing models.

Finally, we may want to expand our notion of structural similarity beyond “high clustering, low diameter, and best fit interpolation parameter between random and scale-free”, in order to better represent the statistical properties of real networks. Jackson’s work focuses primarily on the degree distribution, but we can generalize this idea to that of a *graphlet degree distribution*, as introduced in “Biological network comparison using graphlet degree distribution” (included with annotations of Jackson’s paper)⁵. This modeling strategy has shown itself to be highly suitable for describing protein-protein interaction networks and may have interesting applications in economics and game theory.

⁴According to Table 8.1 in Newman’s *Networks: An Introduction*, 1st edition.

⁵And discussed in chapter five of my thesis project.