Cindy Cook, Matt Denny, Mitch Goist, Timmy Huynh
PLSC 597E / SoDA 502
September 16, 2015

Paper Description and Proposal

1. Overview of the Research Domain

Hierarchy is an important feature of many organizations, such as firms, social clubs, and military units. Formally, we can define a hierarchy as a system where people or groups are ranked according to status or authority. Yet it is difficult to operationalize this definition for measurement and comparison. There has been a great deal of research on power and status in groups and organizations, but most of this research relies on measurements defined over domain specific rankings, such as job titles. At the same time, networks scholars have defined a number of broadly applicable hierarchy metrics based on network structure, but these metrics are not necessarily grounded in meaningful sociological concepts of status and authority. Contrastingly, social theorists like Michael Mann have noted the messiness of society and that a network-oriented perspective of the "sociospatial and organizational model [of a network]" can explicate the "sources of social power," but they have generally not delved into the methodologies through which to fully explore such power dynamics. In this paper, we seek to bring together these two areas of research, and to develop a framework for measuring hierarchy in social networks that is both generally applicable and exhibits a high degree of construct validity.

Having developed a framework for measuring hierarchy in networks, we will then test its internal and external validity. To test the internal validity of such a measure, we will conduct a simulation study. To test the external validity of our measure we will compare measurements across networks that we can theoretically rank by their degree of hierarchy. We will then apply our framework to better understand the implications of hierarchical network structure for organizational performance in a sample of 17 county government organizations.

2. Problem Statement

How do we define and measure hierarchy in (directed) social networks? We need to relate sociological conceptions of hierarchy and power to network measures. In developing this framework we intend to compare analytical and statistical approaches to measurement on both synthetic and real world datasets. In particular there are several key questions we must address.

1. Is an analytical or statistical measure of network hierarchy more appropriate for our goals?
2. Can we capture all or even most salient dimensions of hierarchy as defined in the sociological literature in a single measure?
3. Can our measure be extended to undirected networks?

3. Empirical Research Design Overview

We need to validate our approach on both synthetic and real world data. We should probably stick to directed social networks for now as that is the easier case, and it would be difficult to conceptualize undirected graphs with a distinct hierarchy as opposed to being simply multi-community.

4. Data collection feasibility

We are still working through evaluating a few different datasets to best suit our purposes. However, at present, this is a little difficult because we really want our measure to be theoretically-grounded, but we haven't yet developed a solid theoretical conception for hierarchy. Thus far, theory-wise, the Mann (1986) definition seems closest to the Liu-Driver measures discussed in the Mones et al. (2012) article: i.e., hierarchical networks are those in which the actions of a few nodes are needed to take control of the graph. Another potential definition, also implied, is hierarchy means the mechanisms of collective actions (i.e., the ability of different nodes to connect with one another) hinges on a small number.

Among the network datasets we are exploring, they are already or mostly in usable format. As we're navigating through our theoretical conception of hierarchy in network, we are ruling out the use of the karate club, dolphin, football, etc. datasets because we want to be able to analyze datasets where the networks are more interesting or theoretically-relevant. In this way, there are a couple of systems that might be useful. The first is a network of cooperation among militant groups, which encompasses joint exercises, mergers, and splits among militant groups: http://web.stanford.edu/group/mappingmilitants/cgi-bin/. This may interesting for us for a few reasons: (1) there is no *de jure* hierarchical structure (i.e., no formally-recognized chain of command or sovereignty); (2) militant groups face a classic collective action problem, and thus we can expect the dynamics Mann describes to hold; and (3) most theories of conflict would predict no hierarchy to occur in this system. An interesting system to compare this to would be military actions in Vietnam: http://tinyurl.com/pwofooy. The nodes here would be military units, and the edges are participation in the same battle. Of course, the main issue with this dataset is that it's undirected, which we've noted may be difficult to conceptualize within a hierarchy framework. We're focusing on conflict datasets because many of the theoretical definitions define hierarchy as essentially about outcomes—i.e., the ability of particular nodes to control the actions and behaviors of subordinate nodes.

We may also use manager network data where each organization has a “county manager” who is theoretically in charge of the rest of the actors, providing an opportunity to determine if the methods we employ capture a plausible hierarchical structure. We are still exploring this and other datasets though.

5. Analytics Detail

The analytical portion of the problem will be conducted in R, which is known by all members of the group. We will be using both statistical and mathematical methods of quantifying and/or measuring hierarchy. We will focus on methods that have already been developed, published, and implemented in R, or are can easily be implemented by one of the group members. If time permits, we may try to develop or suggest directions for future development of our own statistical models and/or mathematical measurements. Each member of the group will be responsible for at least one method.

The statistical methods we will be looking into include hierarchical exponential graph models in the R package *hergm*. This package also includes hierarchical stochastic block models. Unlike fitting network data with exponential random graph models (ERGMs), hierarchical ERGMs focus on inducing local dependencies. Next, we will focus on latent space models, which can be fit in R using the *latentnet* package in the *statnet* suite of packages. For both the latent space and ERGM models, Bayesian inferential analysis can be conducted using the *Bergm*, *VBLPCM*, and *lvm4net* packages in R. We note that whenever fitting network data there is always the chance for computational timing and accuracy issues to come up. We have chosen a number of datasets for the purposes of capturing several types of hierarchies, but also so that we may have a few that are easily fit in R. Lastly, we will focus on mathematical measures of hierarchy. These measures primarily stem from graph theory, and can be easily programed by ourselves in R. The measurements include the Global Reach Centrality (GRC), Triangle Transitivity, Kendall's K, and Landau's lambda.

6. Significance

All of our fields including political science, sociology, and statistics approach the concept of hierarchy from different angles. Yet the main goal of any researcher is the same: to accurately theoretically understand and quantify real world phenomena. Without statistical models/mathematical measurements for hierarchy which are theoretically based, and vice versa; theory that can be statistical/mathematically quantified and verified, the conceptual idea of hierarchy cannot be fully understood. We do not suggest that this project will achieve an overreaching theory and methods, but we strive to take the first step. At the very least, we will try to demonstrate the need for a united theory and corresponding methods. As an interdisciplinary team, we are in the unique position to accomplish our goals.

7. Publication Plan

Because our project deals with networks, the most natural publication outlet is the interdisciplinary journal *Social Networks*, which would be a beneficial publication for all group members. As the project progresses, if we find the time to create a new method of measuring for hierarchy(-ies) in networks and compare this new interpretation to prior measures, we may instead aim for the interdisciplinary scientific journal *Science* since we would be pushing for new knowledge. Alternatively, if instead we find that it is more important/interesting for us to contextualize and theorize for the need of measuring hierarchy in networks, we may choose the discipline-specific journal *American Sociological Review* because most of our current ideas and understandings for the operationalization and definitional need of network hierarchy is situated in social theory.