## Poster M-13

# A Computational Framework for Analysis of Dynamic Social Structures

**Authors:**
Tanya Y. Berger-Wolf (*Department of Computer Science, University of Illinois at Chicago*)
Ilya R. Fischhoff (*Department of Ecology & Evolutionary Biology, Princeton University*)
Daniel I. Rubenstein (*Department of Ecology & Evolutionary Biology, Princeton University*)
Jared Saia (*Department of Computer Science, University of New Mexico, Albuquerque*)
Siva Sundaresan (*Department of Ecology & Evolutionary Biology, Princeton University*)

**Short Abstract:** Analysis of many processes in animal populations (disease spread, dominance, reproduction) requires analysis of the network of individual interactions. Until now such analysis was essentially aggregate and static. We show that it may produce incorrect answers and propose a novel conceptual and computational framework for analysis of dynamic interaction networks.

**Long Abstract:**
Understanding why animals do what they do lies at the heart of most studies in animal behavior. Why do some individuals choose to live in groups whereas others live alone? For those that live in groups, why do groups differ in size, composition, and stability? Behavior is the most flexible feature of an animal's phenotype. Behaviors exhibited by collections of individuals determine population processes ranging from movement patterns, to growth rates, to the spread of ideas, genes and diseases, to the nature of social structures themselves. Thus, understanding how population processes emerge from contacts among individuals is a natural outgrowth of understanding why social behavior occurs in the first place. Successfully connecting ecology with behavior and behavior with population processes requires detailed data that until recently has been unavailable to biologists. However, mobile sensor collars and other remote sensing devices are producing this type of data in high quantities and the limitations lie with our ability to manage and analyze these data.
Existing theory, both for analysis of social and other pairwise interaction networks (such as WWW, internet, cellphones), is based on identifying trends that develop over long periods of time. Such a static theory is too crude for determining the linkage between social context and the emergence of fine-grained individual or group behavior. Moreover, we show that static analysis may result in erroneous conclusions when estimating the effects of the population processes. We propose a new conceptual and computational framework that addresses the time component of social interactions explicitly.
In our framework we assume the availability of a time series of observations of the network of interactions of a population. We create an abstract representation of the data as a layered graph. The nodes of the graph correspond to groups of interacting individuals at a fixed timestep. The nodes are arranged in layers of different timesteps. The edges connect between any two groups in different time layers and are weighted by the membership similarity between the groups. The main object of study in this framework is a metagroup. Formally, a metagroup is a sufficiently heavy and sufficiently long path, a sequence of nodes connected by edges. A metagroup is a timeline evolution history of a group of individuals.

Just like any real group, it may change its members over time (depending on the measure of similarity), yet locally remains essentially the same group. Our definition of a metagroup is flexible enough to accommodate a range of social structures, from the “never changing, always together” metagroup to the “meeting every now and then of more or less the same individuals”.

Many questions about the dynamics of social structure, such as persistence and stability of communities and criticality of individuals and interactions, reduce to questions of graph connectivity. For example, while the number of sufficiently long paths in a graph can be exponential in the number of nodes (which makes it impractical to list all the metagroups) we can compute in linear time the number of those paths, their average length, and the longest such path. Those metagroup statistics are indicators about the stability and persistence over time of social structures within a population. Individuals, interactions, and groups may be critical to the social structure of the population if their removal significantly increases or reduces the number of long heavy paths in the metagroup graph. While many questions about population social dynamics can be answered efficiently within our framework many more pose challenging computational problems.

We have applied our framework to the datasets of Grevy zebras (Equus grevyi) from Mpala Research Centre (Kenya) and onagers (Equus hemionus) from India, both collected by the coauthors. The data is from visual scans of the populations typically once a day over periods of several months. Individuals in each observed group (defined by special proximity at a given time) are identified and the GPS coordinates of each group are noted. Our initial results from metagroup analysis coincide with the existing biological understanding of the social structure of the two species: both are fission-fusion species with few short-lived stable communities. Yet, the computational analysis has also revealed some unexpected differences, such as that the communities of onagers are more persistent that previously thought. Future work will utilize GPS sensor collar data for more complete and detailed analysis.

In addition, we have applied the conceptual metagroup framework to a dataset which is considered a benchmark in comparing social networks analysis methods, the Southern women data from 1930s Natchez, Mississippi (Davis et al., 1941). The communities identified by our analysis coincide with those resulting from other (static) social network analysis techniques and have a competence score of .923. Our method also provides a classification of the individuals into the core versus the periphery of their community comparable (but not identical) to the other techniques.

In summary, in this work we demonstrate the need for computational techniques for analysis of dynamic social interaction data in animal populations. Such analysis is both necessitated and enabled by the recent increasing availability of mobile and remote sensor data. We propose a new conceptual and computational framework for analysis of dynamic social networks and present encouraging preliminary results on zebra, onager, and human population data.