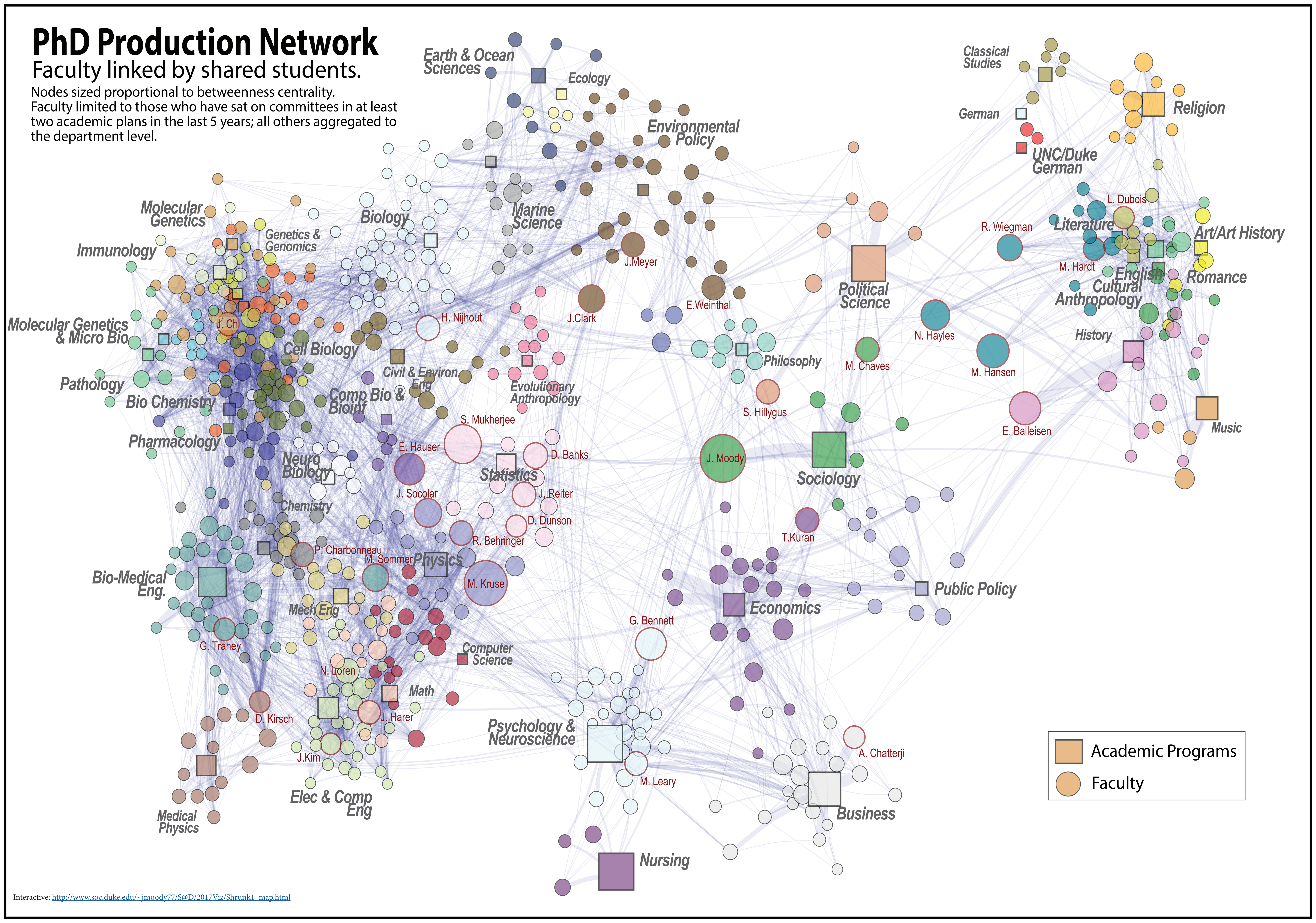
Multiple Sources of Interdisciplinary Training



DNAC Duke Network Analysis Center
DUKE UNIVERSITY

James Moody & The DNAC Lab

Problem

Visualizing interdisciplinary education requires producing a comprehensive representation of graduate education across campus and then layering that representation with interdisciplinarity. The ideal figure(s) should reveal the organization of graduate education while making clear both individual and departmental contributions to interdisciplinary education.

Mapping Graduate Education

The network of faculty linked by shared students nicely captures the structure of graduate education. To highlight faculty who most contribute to interdisciplinary education, we aggregate faculty who sit only on a single program's committees. To highlight the multilevel nature of department and individual bridging, we combine a global distance minimization routine with a program-centered node overlap minimization algorithm and a final pass that places key bridging nodes at the optimal position spanning programs. This three-stage layout routine nicely sorts the university by division, while both highlighting academic programs and the faculty who most often bridge programs.

Measuring Interdisciplinarity

The core problem interdisciplinarity aims to solve is bridging otherwise disconnected academic silos, and we have three ways to think about such bridging: by the structure of the PhD production network, by academic plans, or by scholarly fields. For the network, *betweenness centrality* (Freeman 1977) captures the extent to which faculty connect otherwise disconnected faculty. *Academic program bridging* captures how faculty training crosses PhD programs, while *field bridging* captures how faculty publications cross multiple scholarly fields. While programs are given in the data, we must infer faculty field from faculty publication patterns, since some units employ multiple disciplines. We do so by clustering the Web of Science journal co-citation tables, to generate sets of similarly cited journals, excluding general journals. We then match faculty publications to these clusters to identify a field for each publication. For both programs and fields, interdisciplinarity occurs both within faculty or by committees.

To facilitate multi-level comparisons across the network and content-based measures, we developed a technique that shadows the network layout but tiles nodes into homogeneous blocks. Faculty are represented as tiles that are then colored by the relevant interdisciplinarity score. Scores on the first row represent academic program; those on the second row scholarly field. The first column captures participation, the second column within-person interdisciplinarity and the third committee interdisciplinarity.

Results

The PhD production network is broadly organized by division, with high connectivity within division and low between. The faculty with highest betweenness centrality generally cross divisions. The humanities and interpretive social sciences are fairly well-integrated by program, while the remainder of the social sciences are structurally more insular. Within the natural sciences, seemingly high program interdisciplinarity is less pronounced at the field level, as many faculty from different nominal programs publish in similar outlets.

Technical Tidbits. The faculty sample is limited to 1271 faculty who have served on a PhD committee in the last 5 years and who have data in the provided Scholars@Duke visualization data file. As such, some adjunct, new & emeritus faculty are thus missing. Faculty were assigned to academic programs based on the most common service, with ties given to the mode within their primary appointment organization. We used a fuzzy matching algorithm to link faculty publication journals to the disciplinary cluster file. This resulted in about 80% of faculty publications being matched, with very high levels of missing data in the Humanities and interpretive social sciences and we thus felt it misleading to include them in that analysis. We used a Fruchterman Reingold layout algorithm (implemented in Pajek) to define the base faculty space, then used Kamada-Kawai within academic programs to minimize node overlap. The network tiling procedure is a hill-climbing algorithm seeking to maximize the number of similar neighbors in the neighborhood of each tile. Analysis, cleaning and figure production were done initially in SAS, network figures were produced in Pajek then edited in Illustrator and the poster compiled with InDesign.

References

Batagelj V, Mrvar A . 1998. "Pajek—a program for large network analysis." *Connections* 21(2):47–57.
Freeman, L. C. 1977. "A set of measures of centrality based on betweenness." *Sociometry*, 40:35-41.