

Case Report ■

Using Social Network Analysis within a Department of Biomedical Informatics to Induce a Discussion of Academic Communities of Practice

JACQUELINE MERRILL, RN, MPH, DNSc, GEORGE HRIPCSAK, MD, MS

Abstract In order to assess the mission and strategic direction in an academic department of biomedical informatics, we used social network analysis to identify patterns of common interest among the department's multidisciplinary faculty. Data representing faculty and their self-identified research methods and expertise were analyzed by applying a network modularity algorithm to detect community structure. Three distinct communities of practice emerged: empirical discovery and prediction; human and organizational factors; and information management. This analysis made intuitive sense and served the goal of stimulating discussion from new perspectives. The findings will guide future direction and faculty recruitment efforts. Communities of practice present a novel view of interdisciplinarity in biomedical informatics.

■ *J Am Med Inform Assoc.* 2008;15:780–782. DOI 10.1197/jamia.M2717.

Introduction

Traditional categories that are used to define biomedical informatics center on application areas (e.g., clinical informatics, bioinformatics, public health informatics). While useful and informative, application areas can be an imperfect fit for multidisciplinary academic researchers whose expertise, inquiry and practice cut across categories. A research faculty that is engaged in a range of multi-dimensional activities and that use a range of methods constitutes a complex human system. To understand the relationships within this system we explored the interests of the faculty using social network analysis, an empirical technique for understanding complexity. Our goal was to gain insight to inform an updated mission and to direct future actions such as hiring new faculty.

Social Network Analysis

Social network analysis is grounded in the systematic analysis of empirical data using formal theory organized in mathematical terms.¹ The goal is to capture patterns of human interactions. A network is represented as graph consisting of a finite set of nodes or vertices linked by lines or edges. A "social" network is defined as a group of collaborating (or competing) entities that have some type of relationship and interact within a shared environment often referred to as a community.² Network graphs are notated in the form of an adjacency matrix that allows use of operations from matrix and linear algebra to mathematically define characteristics of the network.³ The

social network can be used to characterize and describe the community structure of its members. Some communities have a densely connected core of members with less central members scattered around the periphery. In others there is fairly uniform distribution of links across people, and still others have obvious divisions into sub-groups or "modules." One approach to quantifying such patterns is a modularity algorithm developed by Newman.⁴ The algorithm reflects the degree to which a network contains community structure by calculating the number of edges that fall within sub-areas of the network minus the expected number in an equivalent network with edges placed at random. The modularity can be either positive or negative on a -1 to $+1$ scale. Positive values greater than 0.3 indicate the likely presence of a community structure.⁵

Case Description

A transition of leadership and a planned recruitment initiative in an academic department of biomedical informatics was the occasion for reviewing the department's mission and strategic direction. As part of that process faculty members were asked to self identify their areas of expertise, research, and practice during the department's annual retreat. The methodological areas and the scope of practice that emerged challenged the faculty to develop a cohesive description of "who we are." To explore the structure of the department we turned to information presented during the retreat which consisted of a single slide prepared by each faculty member summarizing their current research methods and areas of expertise.

Example of Network Analysis Used to Describe Department Structure

We extracted terms or short strings describing the self-identified methods and areas of expertise from the content of the faculty-prepared slides. Similar or closely related items were clustered and de-duplicated within broader terms. For example "evidence-based practice" and "guideline-based care" and "bring scientific evidence into clinical practice" were clustered

Affiliation of the authors: Department of Biomedical Informatics, Columbia University, New York City, NY.

The authors thank the faculty of the Department of Biomedical Informatics at Columbia University who contributed to the development of the concepts presented here.

Correspondence: Dr. Jacqueline Merrill, Department of Biomedical Informatics, Columbia University College of Physicians and Surgeons, 622 West 168th Street, VC-5, New York, NY 10032; e-mail: <jacqueline.merrill@dbmi.columbia.edu>.

Received for review: 01/09/08; accepted for publication: 06/29/08

within decision support. The term “electronic patient charting” was clustered within “electronic health record.” A total of 18 terms were extracted from slides prepared by 25 faculty members. These were arranged in a rectangular adjacency matrix wherein each faculty member who was associated with a term received a score of “1” and all others received a score of “0.” These data were entered into the Organization Risk Analyzer (ORA) computer program developed at Carnegie Mellon University.⁶ We used the Newman modularity algorithm to detect sub-groups in the network. Three sub-groups or “communities” were identified within the department. The modularity score of 0.32 suggested that the groups are relatively distinct, without a high degree of integration. A graphical display of the network colored by Newman groups is shown in Figure 1.

Three communities clustered around activities broadly related to empirical discovery and prediction, human and organizational factors, and information management. Visualization of the data showed clearly not only which people and activities got grouped together, but how three distinct groups emerged. The results resonated strongly with the faculty. Bioinformatics faculty and clinical faculty emerged as a group sharing practice in computational methods such as data mining and knowledge discovery. While there is co-authorship between these faculties, the analysis led us to consider how the two groups may move

the department forward by sharing resources, particularly in the area of curriculum development and training. A series of discussions held via email and during two monthly faculty meetings served to refine the results more inclusively. This process focused the faculty’s energies on defining what the department is, and gleaned insights on future direction and hiring new faculty. The dialogue clarified the relationship of the three communities of practice to the Department’s overarching mission of research, education and service in the context of four recognized areas of biomedical informatics application.⁷ The methods and areas of expertise associated with the communities of practice are displayed in Table 1. These results are already guiding faculty and student recruitment efforts and redesign of the Department’s website.

Discussion

Social network analysis is a method for unraveling complexity. It was applied to gain insight during a period of transition within an academic department of biomedical informatics. When an environment changes, unexamined assumptions may become out-of-date and the actions based on them may become ineffective.⁸ Insight into the complex nature of the department has helped faculty to critique and reconstruct ideas about who we are, and to generate a renewed and shared direction. The process itself assisted in transitioning the depart-

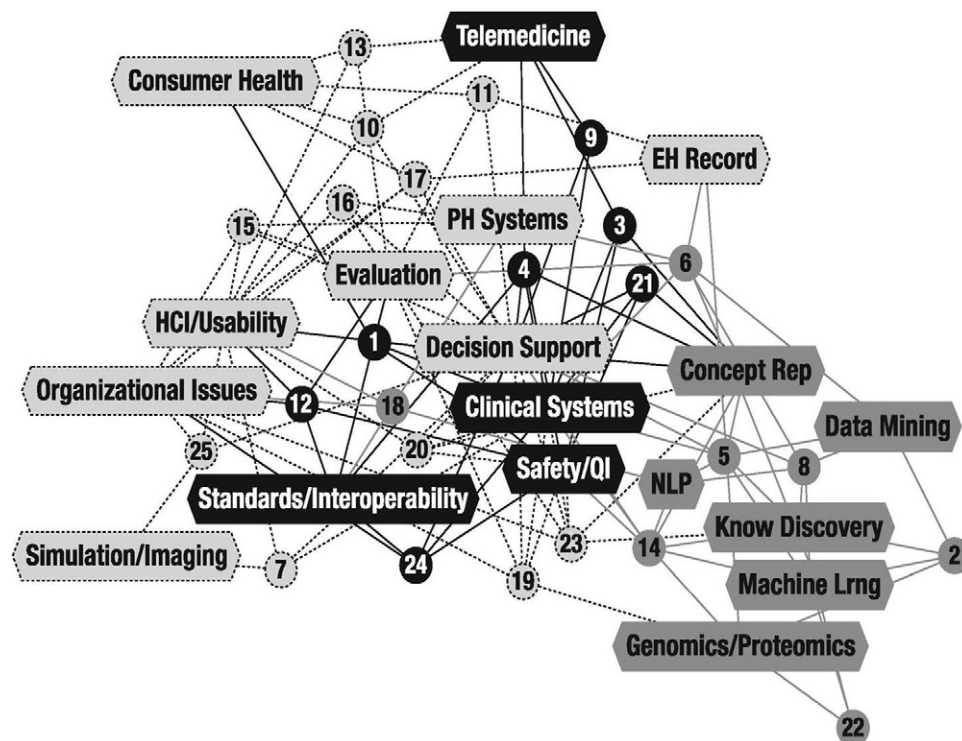


Figure 1. Network Diagram of Faculty Communities of Practice In the network graph each numbered circle is a node representing an individual faculty member; each polygon is a node representing a method or area of expertise. The lines linking circles and oblongs indicate that a faculty member associated him or herself with a particular method or area of expertise when summarizing their research interests for a faculty retreat. The node colors correspond to how faculty and methods or areas of expertise were grouped within one of three communities identified by the Newman modularity algorithm:

Light Grey A community around methods and expertise related to *human and organizational factors*
 Dark Grey A community around methods and expertise related to *empirical discovery and prediction*
 Black A community around methods and expertise related to *information management*

Details on the methods and areas of expertise associated with each community of practice is found in Table 1.

Table 1 ■ Methods and Expertise Associated with Three Communities of Practice with the Corresponding Label on the Nodes in Figure 1

Community of Practice	Method or Area of Expertise Associated with Community	Label in Figure 1
Empirical discovery and prediction	Computational simulation of biological systems focused on genomics and proteomics	Genomics/Proteomics
	Data mining	Data Mining
	Knowledge discovery	Know Discovery
	Natural language processing	NLP
	Concept representation	Concept Rep
Human and organizational factors	Machine learning and intelligent systems	Machine Lrng
	Decision support	Decision Support
	Health technology evaluation	Evaluation
	Consumer health	Consumer Health
	Electronic health records	EH Record
	Public health systems	PH Systems
	Cognition, human computer interface, and usability	HCI/usability
Information Management	Visual simulation of human biological systems	Simulation/Imaging
	Clinical information systems	Clinical Systems
	Safety and quality improvement	QI
	Information standards and interoperability	Standards/Interoperability
	Telemedicine	Telemedicine

ment toward a stronger identity by providing insight into how the scope and range of faculty expertise interacts with the mission of the department. The groups listed in Table 1 are logical and might be grouped similarly if identified by any expert observer. Perhaps the more interesting outcome of the analysis is that the Newman algorithm identified *three* communities of practice as opposed to some other number. That is, an observer operating without the aid of a computational technique might potentially cluster the methods into two, four, or more groups.

The results of a network analysis can be judged by how accurately the network reflects the real world situation, i.e., face validity.⁹ In this case we used self-identified and subjectively categorized research methods and expertise of faculty. Yet when these empirically identified groupings were presented to the departmental faculty for interpretation they just “felt” right, thus satisfying the face validity criteria. The faculty easily came up with labels for the communities of practice represented by the groupings. In the past, the department had loosely organized itself around application areas: clinical, imaging, public health, and bioinformatics. Due in part to external forces like a separate physical location and dedicated funding¹⁰ the bioinformatics faculty constitute a formal division within the department. The other application areas have not formally separated into divisions. The current findings help to explain why this separation has not fully emerged: communities of practice span application areas, preventing any simple separation into divisions.

The groupings produced in this analysis reflect a single point in time. As the department evolves the methods and expertise applied in communities of practice may shift, particularly in areas of cross-cutting practice. For example, electronic health records (EHRs) is grouped within “Human and Organizational Factors” because several faculty were actively working on that aspect of EHRs at the time of the study. Future funding or faculty recruitment may result in more of those within the “Information Management” community applying their energies to EHRs.

Application areas remain an important way to characterize biomedical informatics. The method-oriented communities of

practice identified here supply an alternative “fit” that acknowledges the multi-disciplinary nature of the field. The clustering of methods and expertise associated with these communities of practice have helped faculty in our department understand competencies required for training scientists and have informed curriculum development to improve the quality of our academic training program. The insight gained by identifying communities of practice can help us to meet the challenge of integrating the skills and perspectives of multiple disciplines to build knowledge and to translate that knowledge to solve relevant problems. We studied overlapping areas of methods and expertise rather than actual collaboration among faculty. It would be interesting to compare the two networks to see how communities of practice interact internally and with each other.

References ■

- Freeman LC. The Development of Social Network Analysis: A Study in the Sociology of Science. Vancouver: Booksurge Publishing; 2004.
- Aviv RE, Zippy; Ravid, Gilad; Geva, Aviva Network analysis of knowledge construction in asynchronous learning networks. *J Asynchronous Learning Networks* 2003;7(3):1–23.
- Scott JG, Tallia A, Crosson J, et al. Social Network Analysis as an Analytic Tool for Interaction Patterns in Primary Care Practices. *Ann Fam Med*. 2005 September 1, 2005;3(5):443–8.
- Newman MEJ. Fast algorithm for detecting community structure in networks. *Physics Review* 2004;69.
- Newman MEJ. From the Cover: Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*. 2006 June 6, 2006;103(23):8577–82.
- Carley KM. ORA: Organizational Risk Analyzer. Pittsburgh, PA: Center for Analysis of Social and Organizational Systems; 2007.
- National Library of Medicine. National Library of Medicine Awards \$75 Million for Informatics Research Training. 2006. Available at: http://www.nlm.nih.gov/news/press_releases/award_info_res_training06.html. Accessed September 18, 2008.
- Lane D, Maxfield R. Strategy Under Complexity: Fostering Generative Relationships. *Long Range Planning* 1996;29:215–31.
- Wasserman S, Faust K. 1994. *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.
- Columbia University Center for Computational Biology & Bioinformatics. 2005. Available at: www.c2b2.columbia.edu. Accessed September 18, 2008.