Available online at www.sciencedirect.com
ScienceDirect
Journal of Biomedical Informatics 40 (2007) 422–428
Journal of Biomedical Informatics
www.elsevier.com/locate/yjbin

Description of a method to support public health information management: Organizational network analysis

Jacqueline Merrill [a,*], Suzanne Bakken [a,b], Maxine Rockoff [a], Kristine Gebbie [b], Kathleen M. Carley [c]

[a] Department of Biomedical Informatics, Columbia University, 630 West 168th Street, VC 5, New York, NY 10032, USA
[b] School of Nursing, Department of Biomedical Informatics, Columbia University, USA
[c] School of Computer Science, Carnegie Mellon University, USA

Received 25 April 2006
Available online 5 October 2006

Abstract

In this case study, we describe a method that has potential to provide systematic support for public health information management. Public health agencies depend on specialized information that travels throughout an organization via communication networks among employees. Interactions that occur within these networks are poorly understood and are generally unmanaged. We applied organizational network analysis, a method for studying communication networks, to assess the method's utility to support decision making for public health managers, and to determine what links existed between information use and agency processes. Data on communication links among a health department's staff was obtained via survey with a 93% response rate, and analyzed using Organizational Risk Analyzer (*ORA*) software. The findings described the structure of information flow in the department's communication networks. The analysis succeeded in providing insights into organizational processes which informed public health managers' strategies to address problems and to take advantage of network strengths.

Keywords: Information management; Public health systems; Public health administration; Organizational network analysis

1. Introduction

Public health agencies are complex entities which depend on specialized information which travels throughout the organization via communication networks among employees. Management of the collection, analysis, use, and communication of health related information is considered the most important public health service, undergirding all others [1]. Yet the structure of public health agencies presents considerable challenges to effective information flow [2]. Agencies' individualized structures have evolved non-uniformly as components of state and local governments where public health services are primarily supported by targeted funding streams. Maternal/child health, TB control, immunizations, and HIV/AIDS prevention are some examples of programs that are categorically funded but organized very differently across jurisdictions. The result is information that is not aggregated or accessible accompanied by duplication of effort, information gaps, and strained cooperative working relationships that complicate efforts to accomplish the mission of public health [3,4]. As is true in many organizations, the interactions that occur within internal information networks are poorly understood and therefore are unmanaged processes. Since many public health professionals do not have the skills to make strategic decisions about how information is managed they need methods to help them understand and direct these processes [5,6].

Organizational network analysis is an application of social network analysis, a method that is typically focused on connections between individuals, to an organizational

* Corresponding author. Fax: +1 212 305 0722.
E-mail address: jam119@columbia.edu (J. Merrill).

1532-0464/$ - see front matter
doi:10.1016/j.jbi.2006.09.004

entity. It is a descriptive, empirical research method for mapping and measuring *relationships* between people, groups, and organizations with the resources, knowledge and tasks that are used to perform work. The technique draws upon theories of organizations, networks, and complexity to produce models representing complex organizational interactions that would be infeasible to describe without relational concepts [7]. The resulting insights can help managers understand critical performance factors such as how information diffuses among individuals and influences the speed, quality, and accuracy of organizational decisions [8–10].

The key feature distinguishing network theory and measurement from traditional data analytic methods is use of structural or relational variables analyzed using techniques based on graph theoretic methods [7]. In an organization, networks are comprised of *nodes* that represent agents (human or machine), knowledge, tasks, or resources, and *links* that show relationships between the nodes. Agents have varying degrees of connectivity with other agents through which information and resources flow. Depending on the scale of analysis, an agent may represent an individual, a project team, a division, or an entire organization [11]. These interdependent "node-link" structures, while simple in concept, become related in multifaceted ways as networks grow and develop. Graph theory supplies (i) a vocabulary for denoting social structural properties, (ii) mathematical operations to quantify these properties, and (iii) a method for validating theorems about graphs that can be applied to infer how well they represent social and organizational structures [7,12]. Graph techniques allow networks to be both visually displayed as well as statistically described.

Insights from network analysis can reveal where resources are inadequate for employees to perform their work, or identify problems with how information travels throughout a health agency. The technique can provide empirical data to plan for and justify allocation of resources as well as aid decision-making by revealing links between information networks and process performance. For example, network diagrams and network statistics can show how public health staff are linked to the information they need while working in the field. The results can be related to a specific process outcome for the agency, perhaps how effectively inspections are scheduled or carried out. A local health officer could present such information to a county executive to support a request for hand held data devices.

There is an extensive and exponentially growing body of literature on the application of network methods in the study of organizations [13,14]. Although network analysis is a technique that has proven useful for managing information and improving performance in organizational systems [12,15–17], there is no documentation of the method applied to a public health agency's internal organizational structure. Our purpose here is to describe the method as it was implemented in a health department.

2. Methods

We conducted an organizational network analysis with the goal of assisting public health professionals to identify the value of information management in relation to organizational process. The analysis was guided by three specific aims. The first was to collect relational data on the flow of work-related information from a health department and to produce visual and quantitative models that describe the relationships and flows of information in the department. The second aim was to determine what possible links between information flow and process were suggested by the model. The third aim was to collect feedback from the department's leadership to determine the management value of the analysis and what impact the findings might have on management of the information network.

2.1. Setting and subjects

The setting was a county health department with 156 employees that serve a mixed urban/suburban/semi-rural population of about 300,000. The departmental programs are representative of the range of public health services with adequate numbers of staff to allow network analysis at the program level as well as at the full organizational level. The staff represents a variety of public health titles and programmatic specialization, and is large enough to adequately represent public health workforce issues, such as an aging workforce. The study was a whole-network analysis and was not performed using any sampling technique. Every employee was asked to provide information on his/her relationship with all other employees and analysis included all personnel employed by the department.

2.2. Ethical considerations

Organizational network data are sensitive. Data describe the relationships and position of specific individuals in the organizational network. For the data to be meaningful the researcher must be able to record the link between individuals; thus anonymity in data collection is not possible [7,18,19]. Participants in this study were assured of confidentiality. They were informed that for data analysis and presentation of results, names would be replaced with letters and numbers (e.g. RN1, RN2, etc.), but that it still might be possible to identify them by their position in the network (e.g. if theirs was only such title in a specific program area). The department leaders agreed that individual level data would be shared in a limited and selective manner to ensure protection of confidentiality, and that the focus of any presentations to general management staff would be program level findings, not focused on individuals. Employees were informed that no negative consequences would be associated with participating or declining to participate. The study involved data from public employees about performing their regularly assigned work collected with the cooperation of their employer;

therefore it was exempted from full institutional review for human subjects protections under federal exemption §46.101(b) 5.

2.3. Procedures

Network analysis requires complete data to accurately represent the organization. Employees were informed by the health department's leaders that attendance at data collection sessions was a required work-related activity. Data collection was arranged in conjunction with lunch or coffee breaks as an incentive for participation. Survey completion remained voluntary and a response rate of 93% was achieved, which is sufficient to provide a correlation between true and observed centrality >.90 [20].

The analysis used two sources of survey data to capture network relationships. The first used questions selected to measure work-related relationships and communication among employees, and to elicit patterns in how individuals receive and share information in their routine work [21–23]. The second source was data collected by New York Medical College (NYMC) for separate but related research. These data were generated from a survey on information use that was based on a statement of informatics competencies for public health workers [24]. From the first survey we generated a node set of agents (representing all employees) and four agent × agent relational matrices (representing the relational questions described in Table 1). We converted data from the NYMC survey for secondary analysis by creating nodes sets that were then used to build matrices representing the knowledge, tasks, and resources available to the departments' employees.

On trial analysis the networks produced from the four agent × agent matrices were uniformly dense, an indication of possible expansiveness bias, or error that arises when individuals over report interactions because they exaggerate the characteristics of a relationship and/or have different norms for reporting a relationship [25]. It seemed possible that this group might overstate their ties due to the culture of public health, which associates great value with inclusiveness, collaboration and cooperation. To remedy overstated connectivity the networks were merged [26] and a single matrix was created where positive response to all four questions was entered as a tie ("1"); all else was entered as no tie ("0"). The final organizational network comprised one square agent × agent matrix, and three rectangular matrices: agent × task; agent × knowledge; and agent × resource. The agent node set incorporated attribute data to allow sorting of results by homophily (i.e. common characteristics that can make relationship and communication easier) or proximity. Table 1 displays the data collected for this study and the node sets in which it was used.

2.4. Data analysis using the meta-matrix

Analyses of the prepared matrices were performed using the Organizational Risk Analyzer [ORA] [27], a computational tool that extends network analysis by using a meta-matrix model. The meta-matrix is defined as the networks connecting people, knowledge, resources, and tasks.

Table 1
Survey data collected for organizational network analysis

Survey	Variable	Represents	Node set
Relational			
	Do you receive work-related information from each person listed below?	Who gets information	Agent
	To whom do you give work-related information?	Who gives information	Agent
	Who is important in terms of helping you think about complex problems posed by your work. These may or may not be people that you communicate with on a regular basis.	Who knows what	Agent
	I understand what knowledge and skills this person has. This does not mean I have these skills or knowledge, but I understand what skill and knowledge capacity they possess.	Who does what	Agent
Information use			
	Work title	Homophily	Agent (attribute)
	Experience	Homophily	Agent (attribute)
	Education	Homophily	Agent (attribute)
	Work location	Proximity	Agent (attribute)
	Program	Proximity	Agent (attribute)
	Job level (4 levels)	Cognitive demand	Task
	Information use—relevance of 26 items	Cognitive demand	Task
	Self identified functional roles	Cognitive demand	Task
	Communication with 85 outside agencies	Cognitive demand	Task
	Information use—proficiency using 26 items	Knowledge	Knowledge
	Education level (6 levels)	Knowledge	Knowledge
	Experience level (6 levels)	Knowledge	Knowledge
	Information use—regular use of 26 items	Access	Resource

ORA takes as input one or more matrices in the meta-matrix for an organization, from which it calculates measures that describe the relations among personnel and the tasks, knowledge, and resources they bring to bear on their work [28,29]. *ORA* contains over 100 measures of organizational structure and vulnerability based on work in social networks, operations research, organizational theory, knowledge management, and task management [30].

2.5. Measurements and calculations

A subset of 17 network measures were applied in this study, selected to capture individual prominence and network structure and quality, all of which are described in Table 2. To examine the network positions of individuals we used six measurements. An example is *degree centrality*, that is, the number of direct connections a node (person)

Table 2
Network measures used

Network measures of individual position	
Cognitive demand	Average of graph row vector terms (depending on number of input graphs); measures total amount of effort expended by each agent to do tasks
Degree centrality	Number of direct connections a node has (normalized sum of row and column degrees); indicates how likely a node is to receive what flows through the network
Betweeness centrality	Number of times that connections must pass through a single node to be connected; extent that one person is a broker of indirect connections between all others in network; high scoring individuals can influence and control information flow.
Eigenvector centrality	A variant of degree centrality that shows connections to centrally located nodes. A node connected to many well connected nodes has a high score, but a node connected to many isolates has a low score, even if it has a high degree.
High betweeness centrality, low degree centrality	A node with few direct connections, but if removed from the network will result in a new component. These are boundary spanners that connect their group to others.
Situation awareness between agents	Similarity of actor pairs in social interaction, based on physical distance and socio-demographic data
Network measures of organizational quality	
Diversity knowledge and resources	Distribution of difference in knowledge or resource sharing; measures the degree to which knowledge is equally known or resources are equally accessible. Herfindahl–Hirshman index of market share (economics) applied to column sums of agent × knowledge or agent × resource matrix
Redundancy knowledge assignment, and access	Average number of redundant agents in relation to the knowledge, tasks and resources that we measured. An agent is redundant if there is already an agent that has access to that knowledge item, performs that task, or uses that resource. Column redundancy of agent × knowledge matrix, the agent × task matrix, or the agent × resource matrix
Overall complexity	Density of the meta-matrix. The ratio of the number of edges versus the maximum possible edges for the meta-matrix
Social density	Density of the agent network. The ratio of the number of edges versus the maximum possible edges for the agent network
Shared situation awareness	Average shared situation awareness across agents. The similarity of actor pairs based on social interaction, physical distance, socio-demographic data
Communication Speed, Average	The average shortest path length between node pairs (i,j) where there is a path. If there are no such pairs, then Average Speed is zero. Average inverse closeness centrality for network nodes
Efficiency	The degree to which each component in a network contains the minimum edges possible to keep it connected; degree to which there are efficient communication cycles between agents [31]
Efficiency, global Efficiency, local	Measures efficiency of transporting a piece of information in parallel where all nodes in the network concurrently exchange information (global) or in sequence along nodes in the network (local) [32]. Global measures the closeness of the nodes in the network as the inverse geodesic distances (shortest path) between all node pairs. Local measures the closeness of the nodes in each ego network as the inverse closeness of the ego networks
Network centralization	The centralization of a square network based on total degree centrality of each node. Indicates asymmetry in the distribution of connections, indicates the degree to which communication is centralized around a single agent or small group
Transitivity	The percentage of edge pairs (i,j) (j,k) in the network such that (i,k) is also an edge in the network. A measure of cohesion used to gauge the presence of collaborative groups, a generally positive feature in organizations [33]

has, or how likely that individual is to receive information that flows through the network. Individuals that receive high scores on this measure can be influential because they tend to be well-informed about what is going on in the organization. To examine organizational quality we used 11 measurements. These are related to the degree of cohesion, duplication, redundancy and hence efficiency of the network. For example, *communication speed* identifies the average shortest length of the path between any two nodes and is used to gauge how quickly information moves through a network.

2.6. Preliminary presentation of results

Accurate interpretation of network findings is essential to distinguish patterns from random noise, and to assess the veracity of the network structures generated from the data [34]. Analyses of network ties can be difficult to interpret for two reasons: (1) data are not independent observations and (2) results represent an organization about which the researcher may have little firsthand knowledge to guide interpretation. To conduct a meaningful network analysis, the findings must be interpreted in collaboration with the organization's representatives [16,19]. This step was particularly important for this analysis, as there is no baseline knowledge about the network structure of a local health department.

A preliminary set of ORA visualizations and reports were presented to the department's leaders via web conferencing to collaboratively interpret early findings and stipulate goals for the final analysis of the results. The department's leaders identified three goals for the final network analysis: (1) to capture information about resources and identify where there was inadequate communication, they wanted reports on all programs; (2) to assess potential for change in the programs, they requested information on the distribution of tasks, knowledge, and resources; and (3) they wished to analyze the changes in network structure that might be expected from a planned merger of two divisions.

2.7. Final analysis of network data in ORA

The feedback from the presentation resulted in an analysis plan that included an overall network description, a report on key actors in the network, and a quality report comparing program groups with a detailed section on the planned merger. The findings described the structure of information flow in the department's communication networks. After the final report was distributed, we asked the department leaders to identify ways in which they expected to use the results, and to provide brief feedback on how they thought findings might have an impact. To guide their responses, they were supplied with a list of items from the organizational impact portion of a public health information systems evaluation logic model [35]. These items were: (1) managerial value; (2) changes to organizational processes; (3) redeployment of resources; (4) function changes; and (5) cross program support.

The health department's leaders perceived the results of this analysis to be a useful guide for strategic planning, and found the graphical analyses and other reports useful as a way to think about their agency both in general and at the divisional and programmatic levels. They suggested potential applications of the findings that are within the capacity of public leadership to initiate and which conformed to all five items in the evaluation logic model. These included specific strategies which focused on their goals for the analysis. For example, to address findings that revealed inadequate communication they outlined strategies for orienting new staff across programs areas, to have divisional staff regularly schedule work time in other units, and to institute cross-programmatic briefings. To address potential knowledge loss, to increase shared awareness of resources and tasks, and to take advantage of network strengths such as a well-developed social structure they intend to institute mentoring relationships to ensure transfer of key expertise. They also intend to partner junior staff with more experienced staff to mitigate the effects of staff turnover, a significant potential problem targeted by the analysis.

3. Discussion

Information networks are known to have a strong effect on organizational processes, yet they operate on a different logic of exchange than incentive based markets [36]. This view lends plausibility to organizational network analysis as a evaluation method for public sector agencies that relate their performance primarily to achievement of mission rather than the criteria of financial success used in the for profit world [37]. There is no track record of its usefulness in public health agencies, which operate as part of local government and have comparatively limited capacity for restructuring organizational processes. However, there is increasing public sector awareness that autonomy to manage mission and tasks tends to enhance performance, and increasing recognition that leadership can shape a culture of innovation adaptiveness [38–42]. The skills that public managers need for coping with the constraints of their roles can be improved by accurate surveillance of the task environment. The cognitive accuracy derived from knowing the structure and the central people in a network is political knowledge [43], something with which public agency managers are familiar. Network insights inform management strategies, help managers understand and direct information flows, and supply evidence for planning to improve process and performance. Such insights are equally valuable to managers in the public and the private sectors [44].

We applied organizational network analysis in a pilot study to assess the technique's ability to inform decision making for public health managers by affording insight into organizational process. The study demonstrated that the method has potential utility for public health information management. However, organizational network analysis has been developed primarily to examine organizational structure in the private and military sectors where it has been used successfully to improve performance. Network analysis methods need to be refined to serve the public health domain, and metrics must be researched and developed that can contribute to public health process evaluation. By refining current network analysis techniques and by improving the usability of ORA so that its results can be manipulated more directly by public health managers themselves, ORA output could be used to suggest, or prescribe, preferred organizational structures. Future work will center on improving the accuracy of the public health network data collection, and conducting further analyses to build a comparative database of health department structures to support management decisions and process evaluation.

Acknowledgment

This research was supported by P20 NR 007799 Center for Evidence-based Practice in the Underserved, Columbia University School of Nursing, and NLM N01-LM-1-352.

References

[1] Lasker R, Humphreys B, Braithwaite W. Making a powerful connection: the health of the public and the national information infrastructure. Washington, DC: U.S. Public Health Services, Public Health Data Policy Coordinating Committee; 1995.

[2] O'Carroll PW. Introduction to public health informatics. In: O'Carroll PW, Yasnoff WA, Ward ME, Ripp LH, Martin EL, editors. Public health informatics and information systems. New York: Springer; 2002.

[3] CDC/ATSDR Steering Committee on Public Health Information and Surveillance System Development. Integrating Public Health Information and Surveillance Systems. Atlanta: Centers for Disease Control and Prevention; 1995.

[4] Public Health Foundation. From Silos to Systems: Using Performance Management to Improve the Public's Health. Seattle: Turning Point Performance Management National Excellence Collaborative; 2004.

[5] Mandl KD, Lee TH. Integrating medical informatics and health services research: the need for dual training at the clinical health systems and policy levels. J Am Med Inform Assoc 2002;9(2):127–32.

[6] Ross D. Implication of practicing informatics in the public health sector. J Public Health Manage Prac 2002;8(3):v–vi.

[7] Wasserman S, Faust K. Social network analysis: methods and applications. Cambridge: Cambridge University Press; 1994.

[8] Carley KM. Computational organization science: A new frontier. In: National Academy of Science; 2002, pp. 7257–7262.

[9] Galbraith J. Organization design: an information processing view. Interfaces 1974;4:28–36.

[10] Lumpkin JR, Richards MS. Transforming the public health information infrastructure. Health Aff 2002;21(6):45–56.

[11] Dooley K. Organizational complexity. In: Warner M, editor. International encyclopedia of business and management. London: Thompson Learning; 2002. p. 5013–22.

[12] Kilduff M, Tsai W. Social networks and organizations. London: Sage; 2003.

[13] Borgatti SP, Foster PC. The network paradigm in organizational research: a review and typology. J Manage 2003;29(6):991.

[14] Brass DJ, Galaskiewicz J, Greve HR, Tsai W. Taking stock of networks and organizations: a multilevel perspective. Acad Manage J 2004;47(8):795–817.

[15] Chang M, Harrington JE. Agent-based models of organizations. Handbook of Computational Economics II: Agent-based Computational Economics 2004 [cited2005 February]; Available from: <http://www.econ.jhu.edu/People/Harrington/HCE9-04.pdf>.

[16] Cross R, Parker A. The hidden power of social networks. Boston: Harvard Business School Press; 2004.

[17] Krebs V. An Introduction to Social Network Analysis. 2005 [cited 2005 February]; Available from: <http://www.orgnet.com/sna.html/>.

[18] Borgatti SP, Molina JL. Ethical and strategic issues in organizational social network analysis. J Appl Behav Sci 2003;39(3):337–49.

[19] Scott J. Social network analysis: a handbook. London: Sage Publications; 2000.

[20] Borgatti SP, Carley KM, Krackhardt D. On the robustness of centrality measures under conditions of imperfect data. Social Netw 2006;28(2).

[21] Cross R. Organizational Network Analysis. 2004 [cited 2005 February]; website with illlustrations]. Available from: <http://www.robcross.org/sna.htm/>.

[22] Krackhardt D, Hanson J. Informal networks: the company behind the chart. Harvard Business Review 1993;71(4):104–12.

[23] Newman MEJ. The structure and function of complex networks. 2003 [cited 2005 March]; Available from: <http://www-personal.umich.edu/~mejn/courses/2005/cscs535/review.pdf/>.

[24] O'Carroll PW. Public Health Informatics Competencies Working Group. Informatics Competencies for Public Health Professionals. Seattle: University of Washington, Northwest Center for Public Health Practice; 2002.

[25] Feld SL, Carter WC. Detecting measurement bias in respondent reports of personal networks. Social Netw. 2002;24:365–83.

[26] Kumbasar E, Romney AK, Batche lder WH. Systematic biases in social perception. Am J Sociol 1994;100:477–505.

[27] Carley KM, Reminga J. ORA: Organization Risk Analyzer. Pittsburgh, PA: Carnegie Mellon University, Center for Computational Analysis of Social and Organizational Systems (CASOS); 2004.

[28] Carley KM, Kamneva NY. A Network Optimization Approach for Improving Organizational Design. Pittsburgh, PA: Carnegie Mellon University, Center for Computational Analysis of Social and Organizational Systems; 2004.

[29] Krackhardt D, Carley KM.A PCANS Model of Structure in Organizations. In: International Symposium on Command and Control Research and Technology; 1998; Monterey, CA; 1998.

[30] Reminga J, Carley KM. ORA Measures Document. 2005 [cited October 2005]; Available from: <http://www.casos.cs.cmu.edu.>.

[31] Krackhardt D. Graph theoretical dimensions of informal organizations. In: Carley KM, editor. Computational organization theory. London: Lawrence Erlbaum Associates, Inc.; 1994. p. 89–111.

[32] Latora V, Marchiori M. Efficient behavior of small-world networks. Phys Rev Lett 2001;87(19):198701-1-4.

[33] Schank T, Wagner D. Approximating Clustering-Coefficient and Transitivity. 2004 [cited 2005 October]; Available from: <http://www.ubka.uni-karlsruhe.de/cgi-bin/psview?document=ira/2004/9/>.

[34] Breiger RL. Emergent Themes in Social Network Analysis: Results, Challenges and Opportunities. In: Workshop on Dynamic Social Network Modeling and Analysis; Washington, DC: National Academy of Sciences; 2003.

[35] Association of Public Health Laboratories, Public Health Informatics Institute. Toward Measuring Value: A evaluation framework for public health information systems. 2005 [cited 2005 October]; Available from: <http://www.phii.org/Files/FrameworkBrief.pdf/>.

[36] Uzzi B. The sources and consequences of embeddedness for the economic performance of organizations: the network effect. Am Sociol Rev 1996;61(4):674–99.

[37] Provan KG, Milward HB. Do networks really work? A framework for evaluating public-sector organizational networks. Adm Rev 2001;61(4):414–21.
[38] Government Performance and Results Act, 1993. In.
[39] General Accounting Office. Reports on the Government Performance and Results Act. 2003 [cited 2005 November]; Available from: <http://www.gao.gov/new.items/gpra/gpra.htm/>.
[40] Gore A. From red tape to results: Creating a government that costs less. Report of the National Performance Review. In: Washington, D.C.: U.S. Government Printing Office; 1993.
[41] Rainey HG. Comparing public and private organizations: empirical research and the power of the a priori. J Public Adm Res Theory 2000;10:447–69.
[42] Rainey HG, Steinbauer P. Galloping elephants: developing elements of a theory of effective government organizations. J Public Adm Res Theory 1999;9(1):1–32.
[43] Krackhardt D. Assessing the political landscape: structure, cognition, and power in organizations. Adm Sci Q 1990;35:342–69.
[44] O'Toole LJJ. Treating networks seriously: practical and research-based agendas in public administration. Public Adm Rev 1997;57(1):45–52.