Chapter 13
Mining Organizations' Networks:
Multi-Level Approach[1]

James A. Danowski
University of Illinois at Chicago, USA

ABSTRACT

This chapter presents six examples of organization-related social network mining: 1) interorganizational and sentiment networks in the Deepwater BP Oil Spill events, 2) intraorganizational interdepartmental networks in the Savannah College of Art and Design (SCAD), 3) who-to-whom email networks across the organizational hierarchy the Ford Motor Company's automotive engineering innovation: "Sync® w/ MyFord Touch", 4) networks of selected individuals who left that organization, 5) semantic associations across email for a corporate innovation in that organization, and 6) assessment of sentiment across its email for innovations over time. These examples are discussed in terms of motivations, methods, implications, and applications.

OVERVIEW

When you think of social network analysis you probably visualize individuals as nodes. This is quite natural, given the "social" aspect, and more specifically because the origins of social network analysis, going back some 89 years (Freeman, 1996), are in the relations among individuals. Nevertheless, 'social' is also considered at levels of analysis in which the focal nodes are groups, subunits of organizations, organizations, or more macro-level human systems. In this chapter I center on organizational social network analysis, focusing on interorganizational, organizational, departmental, and also individuals as they communicate with organizations.

After this overview I will discuss why such a focus is desirable in light of the literature. First, however, let me point out that this chapter also includes mining for and automatic identification of an organization's networks at several levels based on textual elements of documents available on the web such as news story databases, blogs, reports, and other electronic text content, and also

DOI: 10.4018/978-1-61350-513-7.ch013

available from internal organization documents, such as email, or other kinds of electronic text accessible in real time or as archived internally or in the clouds.

Imbedded in electronic texts are layers of social networks that can be unpeeled with the miner's tools. The highest-order type of network mining I focus on is identification of *interorganizational* networks. This has been an area of research that previously used manual procedures other than mining (Scott, 1988, 1991, 2000; Galaskiewicz, & Shatin, 1981; Galaskiewicz, 1985; Mizruchi, 1996). Here I use an automated procedure based on the co-appearance of organizations across a corpus of web documents. In particular, each time a focal organization appears together with another organization within a proximity window in a document, the pair of organizations is automatically counted. Taking the aggregated collection of pairs, I network analyze them. This enables indexing the organizations' positions in the network and computation of various network structure measures common to social network analysis (SNA). Moreover, by slicing the collection of documents into time segments one can analyze the network structure as a time series. Time-sequenced associations are one of the necessary conditions for establishing possible causal relationships among variables. For example, we could also measure the sentiment expressed in the documents about the organizations and identify any synchronous or lagged associations between sentiment and network structure (Danowski & Cepela, 2010). As an example of this, Noah Cepela and I measured the organizational networks among U.S. presidents' cabinets over time, from Nixon through G.W. Bush, automatically indexing co-appearance of cabinet members in documents. We examined the time-lagged relationships between the presidents' centrality in the administrative network and the link between news sentiment and job approval as measured by the Gallup polls. We tailored the time slicing to the frequency of the Gallup polls for each presidency. This shows how one can build some of the necessary conditions for causality evidence: identify a sequence of networks, index its attributes, and add measures of other attributes of the organizational actors and contextual factors of theoretical interest. What remains is ruling out rival explanations for observed time-ordered associations.

In this chapter my first several examples illustrate such a new approach to interorganizational social network analysis and data mining. I compile a list of organizations of interest and search across large text corpora for the co-appearance of pairs of organizations in news documents, blogs, reports, and related venues. In the first example I mine for an interorganizational network. I also slice time segments across the larger mining time frame to enable time-series analysis of these network structures. The example examines the interorganizational networks associated with coverage of the Deepwater BP Gulf Oil Spill of 2010. I then examine the relationship between news story sentiment about the most central organization, BP, and its position in the network overtime using a new network-based sentiment measuring approach, based on average shortest paths from BP to each of several thousand possible sentiment words, generating theoretically interesting findings about the sequencing of sentiment and centrality.

A second kind of organizational network mining is for the departments within an organization that co-appear across news stories and other text documents. For example, consider the departments within a university. In many cases the departments are proxies for disciplines. By identifying each pair of departments co-mentioned across a corpus of news and/or blog content about the university one can automatically map the representation of the collaborative network of the university's departments over time in the mining corpora. My example is from the Savannah College of Art and Design (SCAD) that was under accreditation review in 2009 and needed evidence of collaborative networks across departments, surrogates for disciplines.