Words and Learning Networks

Caroline Haythornthwaite
Director and Professor
School of Library, Archival and
Information Studies
University of British Columbia
Suite 470, 1961 East Mall
Vancouver, BC V6T 1Z1
c.haythorn@ubc.ca

ABSTRACT
This brief paper outlines research and interests at the nexus of text analysis and social network analysis, with a focus on the application to questions of learning networks.

Categories and Subject Descriptors
J.4 SOCIAL AND BEHAVIORAL SCIENCES

General Terms
Design, Human Factors.

Keywords
Text analysis, social network analysis, learning

1. INTRODUCTION
One of the opportunities presented by the combination of text analysis and network analysis is the potential to ‘see’ learning through the analysis of online conversational transcripts. Whether in open conversations on the web (in discussion lists, virtual communities, or wiki talk pages) or in closed learning spaces (e.g., learning management systems) transcripts of online conversations provide a wealth of information on the process of idea formation, knowledge exchange, argumentation, and entry into communities of practice (Haythornthwaite & de Laat, 2011). Social network analysis provides ways of approaching and modeling online crowds and communities (Haythornthwaite, 2007, 2009) with the combination of text analysis promoting a more substantive understanding of relations and ties (Gruzd & Haythornthwaite, 2011).

2. RESEARCH
My research in general examines how online means of communication and interaction support social relationships online. In examining social networks, I have looked at who talks to whom about what, and via which media. To date, most of this work has determined the “about what” part a priori, framing the social network question in a way to elicit information about the nature of the relationships, e.g., “Who have you supported in a major or minor crisis?” Other work has pulled the relationship qualitatively, examining answers to a questions such as “Who did you learn from?” following up with “What kinds of things did you learn from them?” While these approaches are still viable, the appeal of automated text analysis combined with social network analysis is the way it opens up exploration of much larger datasets, and more networks. Thus, more recently my work has turned to automated methods of managing data collection, text analysis and network analysis.

The research to date has explored some beginning aspects of combining text analysis with social network analysis, with the expectation of applying this to learning processes. Work with Anatoliy Gruzd entailed deriving a means of using text analysis to derive better social network data from online, threaded conversations (Gruzd & Haythornthwaite, 2008). The method involves using text analysis to identify greeting and sign-off text within message content that indicated who was being addressed. The procedures used probability models to determine if in-text names were actually references to authors rather than participants to increase accuracy of network name detection. Gruzd (2009) was able to show that across six online learning classes these ‘name networks’ revealed 33–40 per cent more information about social ties compared to ‘chain networks’ that relied only on ‘who posted after whom’ data. Gruzd’s ‘Netlytic’ environment (http://netlytic.org/) makes this technique available to others, and provides a way to prepare and clean data for analysis.

A different approach was taken in the examination of interactivity in eight online classes (31-54 participants; 1205 to 2156 postings from the class-wide discussion boards; Haythornthwaite & Gruzd, 2012). To explore social networks and threading activity, first all postings within a class were sorted into threads according to the subject line. A tie was taken to exist between two postings if one posted directly after the other in the sequence of this thread. The purpose was to explore the range and extent of interactivity in the classes as demonstrated by thread number, length and participant engagement. While the text analysis was a relatively unsophisticated exact matching, the results were useful for gaining an understanding of threading activity, particularly in showing similarities and differences across the eight classes. While the social networks analysis also has limitations, they provided a beginning look at interaction patterns, showing many one-time pairings, but few repeat pairings of post and response. However, a few cliques of repeat pairings were present suggesting follow-up is to look for the emergence of networks of those watching and following the postings of others in these classes.

3. NEXT STEPS
At this stage, I am interested in learning more about text analytics and determining useful ways of combining text and network analysis with questions about learning and teaching. As attention to cyberinfrastructure turns to matters of student evaluation,
outcomes assessment, and educational success, it is important that this does not become reduced to assessments that address quantity without a view to quality. It is in text analysis that I see a way to increase our understanding of quality interaction. This approach is in concert with the growing awareness that the next generation of analytics tools can only be turned to addressing learning if the groundwork in techniques and studies are there to inform their use (see the emerging area of learning analytics; http://www.solaresearch.org/). I see the combination of text analysis and social network analysis is able to provide a greater understanding of learning as a relation that connects people, but also addressing the composite of interactions that support a learning tie (e.g., trust, social support, reciprocity; Haythornthwaite, 2006). At a network level, such analyses can provide an understanding of learning as the outcome of such relations and ties, e.g., as the social capital of networks (Haythornthwaite & de Laat, 2011).

4. REFERENCES


