Article

A sociogram is worth a thousand words: proposing a method for the visual analysis of narrative data

Qualitative Research
1–18

DOI: 10.1177/1468794116682823
journals.sagepub.com/home/qrj
SAGE

Damien Contandriopoulos
University of Montreal, Canada

Catherine Larouche
McGill University, Canada

Mylaine Breton
University of Sherbrooke, Canada

Astrid Brousselle
University of Sherbrooke, Canada

Abstract

This article proposes an innovative method for the visual analysis of narrative data that involves three steps: transforming narrative data into relational data, creating two-mode networks displayed with graph optimization algorithms derived from social network analysis (SNA), and visually analyzing sociograms. We argue that understanding how actors and their opinions constitute a network-like structure opens up promising avenues for interpreting data. This approach provides powerful data visualization that facilitates inductive identification of the underlying structure of narrative data. It also reveals the complexities of the links between differently positioned actors in a structure that a personal attribute-based analytical method might overlook. Lastly, it can be productively combined with other quantitative and qualitative methods to make sense of narrative data.

Keywords

healthcare, narrative data, Quebec, social network analysis, sociogram, visual analysis

Corresponding author:
Damien Contandriopoulos, Faculty of Nursing, University of Montreal, C.P. 6128 succ. Centre-Ville, Montreal, Quebec, H3C 3J7, Canada.
Email: damien.contandriopoulos@umontreal.ca

Introduction

This article proposes and illustrates an innovative method for the visual analysis of narrative data that involves three steps: transforming narrative data into relational data, creating two-mode networks displayed with graph optimization algorithms derived from social network analysis (SNA), and visually analyzing sociograms.

This method was developed and pilot tested in the context of a research project focused on identifying policy options to improve the performance of the public healthcare delivery system in Quebec (Canada). The complexity of the content and structure of the interviews conducted in this project defeated the usefulness of direct transcript coding and personal attribute-based thematic analysis and prompted the development of the method presented here. Ultimately, examining the network-like structure constituted by actors and their opinions offered more promising ways of interpreting the data. This approach presented three main advantages. It provided powerful data visualization that facilitated inductive identification of the underlying data structure. It revealed the complexities of the links among differently positioned actors in a structure, which a personal attribute-based analytical method would have overlooked. Finally, it ensured effective protection of the confidentiality of information provided by well-known stakeholders in the healthcare system without jeopardizing meaningful interpretation of the research results.

The method we developed is grounded in social network analysis (SNA), an approach focused on understanding the structure of the relations connecting different elements. Although it has a long tradition in social sciences, its mainstream acceptance is recent. This surge in interest stems mostly from the development of powerful and user-friendly software, as well as from the discovery of new applications in genetics, molecular analysis, and biochemistry (Heath et al., 2009; Scott, 2000; Scott and Carrington, 2011).

We begin with a brief summary of the specificity, origins, and evolution of SNA tools, followed by a discussion of how the reliance on relational analysis differentiates SNA from other paradigms used in qualitative and quantitative analysis. We then present the narrative data analysis method we developed, illustrating it with a case study analyzing policy options for reforming Quebec's healthcare system. Lastly, we outline the originality, potential, and applicability of using SNA-based methods to analyze narrative data and conclude by arguing the value of this specific use of SNA in social sciences.

Social network analysis within the quantitative/qualitative debates

At its core, SNA is the analysis of the relations (ties) that link elements (nodes) in a network. It studies network structures or, in other words, the patterns that unite different elements. Being essentially a generic conceptualization of network structure, SNA is not restricted to the analysis of social relations. It has deep roots in various social sciences disciplines but has also followed its own paths of development in a variety of other disciplines, from early 18th-century graph theory to the recent modeling of real-world networks such as the Internet, infrastructures, and biological and social networks (Gastner and Newman, 2006; Knoke and Yang, 2008; Newman et al., 2006). It is beyond the scope of this article to trace a comprehensive history of this field, and several resources provide extensive information on this matter (Freeman, 2004; Kadushin, 2012; Knox et al., 2006; Prell, 2012; Scott and Carrington, 2011; Wasserman and Faust, 1994). Let us simply note here that the main interest of social scientists, in their early experimentations with SNA, was to find new methods to gather, arrange, and make sense of data on social relations (Crossley, 2010). It was also a way to conceptualize social structures as 'networks of actually existing relations' (Radcliffe-Brown, 1940: 2) and depart from conceptual approaches focused on individual entities and personal attributes of actors (Knoke and Yang, 2008; Lorrain and White, 1971; Marin and Wellman, 2011; Scott, 2000; Scott and Carrington, 2011).

With the recent availability of powerful and user-friendly computer programs, this field has witnessed a rapid evolution, such that SNA has become a truly transdisciplinary, widely known and valued tool (Heath et al., 2009; Knoke and Yang, 2008; Palla et al., 2005; Scott, 2000). Recent literature indicates, however, that in social sciences, it is its 'quantitative' dimension that has progressed the most in recent years, despite SNA's long engagement with qualitative methods (e.g. Barnes, 1954; Bott, 1955; Heath et al., 2009; Mitchell, 1969). This situation has fueled new debates on the nature of SNA and the place given to qualitative approaches, with many advocating for stronger integration of qualitative SNA (Heath et al., 2009; Knox et al., 2006) and better recognition of the value of mixed-method approaches in SNA (Bellotti, 2014; Crossley, 2010; Domínguez and Hollstein, 2014; Edwards, 2010; Mische, 2003).

In relation with those debates, we would argue that SNA does not fit squarely into a quantitative/qualitative methods typology. While generally it makes sense to distinguish between qualitative and quantitative methods, in practice this distinction is often—and in our view wrongly—extended to the nature of the data itself. We believe it to be both erroneous and counterproductive to label data with a method-derived typology. We suggest instead that data should be described as either narrative or numerical. To convey meaning, narrative data relies on language (Bourdieu, 1982; Chouliaraki and Fairclough, 1999; Saussure et al., 1955) and numerical data relies on numbers.

What is crucial to understand is that the nature of the data does not *per se* dictate the nature of the methods (Morgan, 1993). For example, there are many methods to analyze narrative data quantitatively (Franzosi, 1994). Word occurrence counts, frequency of passive voice relative to given subjects, and average proximity in number of words between terms or concepts in sentences are all straightforward quantitative indicators to make sense of narrative data (Contandriopoulos et al., 2012; Hsieh and Shannon, 2005; Krippendorff, 2012; Riff et al., 2014; Tausczik and Pennebaker, 2010). Conversely, the same narrative data could be interpreted qualitatively, either directly or through codification, to derive implicit or explicit meaning (Blommaert and Bulcaen, 2000; Fairclough, 2013; Hardy et al., 2000; Hsieh and Shannon, 2005; Mauws and Phillips, 1995; Van Dijk, 2011; Wodak and Meyer, 2009). Likewise, a given set of numerical data could be plotted in descriptive graphs and interpreted qualitatively or submitted to quantitative probabilistic testing of the plausibility of hypothesis. Our point here is simply that the nature of the data does not dictate the choice of either quantitative or qualitative methods. As Small writes, the terms 'qualitative' and 'quantitative' do not represent all-encompassing categories of research and should be used with specific reference to different kinds of data, types of data collection or analytical methods (Small, 2011; see also Crossley, 2010).

This distinction is important for our argument, as SNA is intrinsically neither a qualitative nor a quantitative approach. It can be qualitative, quantitative, or both, depending on the nature of the data collected and the methodological and analytical choices and preferences of the researcher. Relational or network-based data can be analyzed quantitatively or qualitatively, and, in like manner, both numerical and narrative data can be formatted in ways that allow relational analysis.

The data we used as a starting point are narrative and, contrary to the experience of many authors (Bidart and Lavenu, 2005; Heath et al., 2009; Hollstein, 2011; Keim et al., 2009; Mckether et al., 2009), do not *a priori* contain direct relational information. Only a few research projects have experimented with this somewhat atypical use of SNA. For example, Bearman and Stovel transformed written autobiographical accounts of becoming and being a Nazi into relational data by using elements of the narrative as nodes and narrative clauses as ties between the nodes. The resulting 'narrative network' provided new insights into the process of identity formation (Bearman and Stovel, 2000; see also Smith, 2007). In another example, Light transformed American presidential inaugural speeches into relational data (speech-to-speech and word-to-word matrices) to analyze the similarities and differences in presidents' opinions and rhetorical strategies (Light, 2014). Still, this remains a relatively narrow field.

In the next section we propose a method for codifying corpora of complex narrative data into matrices that can be used to conduct relational analysis. This method is based on transforming narrative data into relational data and then using graph optimization algorithms derived from SNA to draw sociograms from this data. We conclude with suggestions on the interpretation of results.

One- and two-mode networks

The simplest way to organize the structural information used in SNA is in a matrix, where columns and rows represent nodes and where the entries represent ties. When the columns and rows of the matrix are identical, the corresponding network is called a one-mode network. This is the most common situation in SNA (Borgatti and Everett, 1997). As an example, a matrix could represent which classmates play with each other outside the classroom in the following way:

Our analysis is based instead on a two-mode network, in which rows and columns refer to different kinds of entities (Borgatti and Everett, 1997; Borgatti et al., 2013). In such networks, rows in the matrix could represent individuals, whereas columns could be social events, such as cultural celebrations, sporting competitions, town hall meetings, etc. In the two-mode network example above, the matrix connects two different elements: people and leisure activities.

The entries for each person correspond to their participation in each activity. As the method we propose is based on two-mode network analysis, it is crucial that this distinction be made clear. In the one-mode example above, individuals have interpersonal connections. In the two-mode example, what connects individuals is their participation in the same activity.

Table 1. Example of one-mode and two-mode networks.

One-mode	Andrew	Adam	Addison	Abigail	Adrian
Andrew		X		X	
Adam	X			X	X
Addison				X	
Abigail	X	X	X		
Adrian		X			
Two-mode	Baseball team	Reading club	Athletics	Judo	World of Warcraft
Andrew	X		X	X	
Adam		X			X
Addison				X	
Abigail		X	X	X	X
Adrian		X		X	

Illustration of the method

The project: Collecting stakeholders' views about problems and solutions to improve performance and ensure sustainability of Quebec's healthcare system

The method showcased here was developed in the context of a research project focused on identifying policy options to improve the performance of the public healthcare delivery system in Quebec (Canada). One component of this project was based on semi-structured interviews with informants who occupied, or had occupied, key positions within organizations, potentially influencing the evolution of the health system (deans of health profession faculties, presidents of professional associations, CEOs of large institutions, etc.). The interviews were loosely structured around four questions: 1) What are the strengths of the current healthcare delivery system? 2) What are the main challenges and problems facing the system? 3) What solutions are needed to tackle those challenges and improve the system's performance? 4) Who are the most powerful actors and interest groups able to shape policy-making in the healthcare system? The team conducted 31 interviews lasting from 45 minutes to two hours.

The research team faced many unexpected challenges in trying to make sense of this body of data. The main stumbling blocks were that the content and structure of the interviews varied considerably from one informant to the other, and that the data were composed of sophisticated arguments in which the meanings of many statements depended on previous assertions and for which the causal links were spread over many pages of transcripts. Classical methods such as direct transcript coding were relatively unhelpful, since coding whole pages did very little to make sense of the content. These challenges underscored the need for data systematization and standardization methods that could lead to more productive forms of analytical reasoning. Relational analysis offered this possibility, prompting the team to develop the method proposed here.

Schemata-based coding

The first step involved working with summaries of core ideas and themes found in the interviews to produce a dataset that was more amenable to relational analysis. Operationally, the process involved two different researchers independently summarizing each interview. Summaries were then compared, discussed, and combined into a single text. The latter was used inductively to build a list of themes or schemata (DiMaggio, 1997) representing the essence of each distinct complex idea, proposition, or statement put forward by each informant. Each schema was summarized in a few words to facilitate coding but was related to a longer definition in a codebook. Schemata were organized according to the four topics covered by the interview questions: strengths, problems, solutions, and influential groups. Two researchers worked independently on developing schemata inductively while at the same time coding each transcript. The coding difference rate (the number of schemata-based codes that were used by only one of the two researchers over the total number of codes) varied from 3% to 26% (average 15%) for the 31 transcripts. Schemata were then discussed in team meetings to identify duplicate ideas, overlapping concepts, or too-narrowly defined codes. Some codes were then merged or removed and discrepancies in interpretation were resolved. Of the 125 schemata-based codes initially proposed, 106 remained at the end of the process: nine strengths, 41 problems, 46 solutions, and 10 influential groups. Once the codes were finalized, all summaries were reviewed to ensure coding was consistent throughout the corpus.

These schemata obtained from the 31 interviews were then incorporated into an MS Excel matrix linking each informant (columns) to all the schemata (rows) corresponding to that particular interview. Each schema was assigned a separate row, and an ‘x’ was placed in the column of each informant who had raised this schema/theme in their interview. To store informant and schema characteristics, we used lines and columns labels with concatenated names built according to a standardized syntax. For example, for problem-based schemata we used the syntax *P–sequential number–descriptive name*, for solution-based schemata, *S–sequential number–descriptive name*, and for informants, *I–sequential number–professional role*. Including node characteristics in the labels is an easy way to retain this information when the data is exported to SNA software. The end product of this process was a schemata (rows)-to-informants (columns) matrix summarizing our 31 interviews and thereby—and this is a key contribution of our method—transforming the nature of our data from narrative to relational. In the process, as we discuss below, some information was lost, but by using this synthesis process based on complex schemata, we were able to retain the sophistication of informants’ ideas, preferences, and opinions.

A key feature of our proposed method is that a matrix such as the one produced here is readily importable into any SNA software to be transformed into a visual form without information loss. Organizing data into matrices facilitates their effective mathematical manipulation, but is a less effective way for humans to grasp information than is a visual representation. Small datasets can easily and quite effectively be drawn by hand—as was done by the pioneers of SNA—but larger datasets, such as the case example presented here, benefit greatly from the computational power of SNA software. It offers the options of using visual optimization algorithms that facilitate the reading of complex and dense networks and of plotting numerous quantitative indicators of node, tie, and network characteristics on the sociograms. For this research, we used primarily Cytoscape 3.1.0, which is multi-platform, free, open source, and relatively user-friendly, while also offering sophisticated customization options for sociogram visualization. However, some of our matrix operations also necessitated the use of more elaborate software, and for those we relied on Ucinet 6. Overall, the method proposed here can very likely be applied using any standard SNA software or a combination of software.

Once we imported the schema-to-informant matrix into the software, we obtained a visual representation of our two-mode network comprised of 137 nodes and 861 links. Each node represented an informant or a schema, and the links tied each informant node with each of the schema nodes corresponding to that informant's interview.

Producing and interpreting optimized sociograms

Force-directed graph optimization. To optimize the visual display of the relational data imported into the SNA software, we used 'force-directed' sociogram optimization algorithms (Cline et al., 2007; Shannon et al., 2003). Force-directed sociograms are optimized from the principle that nodes are mutually repulsive to each other, while edges constitute attractive forces. Force-directed algorithms ensure that repulsive and attractive forces are balanced so that interconnected nodes are closer to each other. In resulting sociograms, highly interconnected nodes will be pushed to the center of the network and clusters of interconnected nodes (cliques) will be visible. There are numerous force-directed algorithms, some of which allow the optimization parameters to be customized. The sociograms presented here are based on the Fruchterman–Reingold algorithm, as the mathematical bases of this model are easily available (Fruchterman and Reingold, 1991). For validation purposes we compared these with sociograms obtained from the various force-directed algorithms available in the two software programs used and found no major differences in the interpretation of the data. Nevertheless, even though force-directed algorithms all tend to produce similar patterns, one drawback of relying on graph optimization algorithms for analyzing data is that those algorithms were developed with the goal of producing visually satisfying graphs, and the optimization parameters they are based on are somewhat subjective. To further test the reliability of the visual analysis method proposed here, we used another visualization technique based on completely different mathematical foundations, as we discuss below.

The visual interpretation of graphs. Our choices of optimization algorithm and ways of plotting indicators on the graph led us down two interesting avenues for visually analyzing how consensual a given schema was among informants. First, since node sizes are mapped proportionally to degree (number of ties connected to one node), larger schemata-based node sizes symbolize greater prevalence of schemas in informants' views (salience). Second, the use of a force-directed graph optimization algorithm means that the more a schemata-based node is in a balanced position between clusters of informant-based nodes, the more consensual it is among those informants. Conversely, nodes that are more interconnected with a sub-group of informants than with the average will be positioned closer to that sub-group. Nodes that are poorly interconnected will be isolated from the others and positioned closer to the periphery. The most interesting feature of this visualization approach is that nodes sharing many links will be positioned together, far from the nodes with which they share few links. This clustering effect allows visual assessment of convergence between informants, between schemata, and between informants and schemata. Obviously, visual analysis is somewhat subjective and imperfect, but the nature of two-mode data prevents the use of most clustering algorithms, as there are no direct connections both between informants and between schemata.

One-mode transformation. To visually assess the similarities between schemata and between informants' opinions, and to validate that the force-directed algorithms could indeed be interpreted in the way we propose, we transformed our two-mode matrix into one-mode matrices.[1] Two-mode matrices can be transformed into two sets of one-mode matrices (Borgatti, 2009a; Borgatti, 2009b). The first one-mode matrix will be obtained by computing the number of times each pair of rows occurs in the same column, and the second by computing the number of times each pair of columns occur in the same row. The Ucinet 6 software automates this process and can compute Jaccard similarity coefficients instead of plain occurrence counts for those matrices. The resulting one-mode matrices of similarity coefficients can be plotted using metric multidimensional scaling (MDS). This will produce graphs where schemata generally discussed together will be positioned near each other and where informants discussing similar schemata will also be positioned close together. MDS graphs are more reliable than force-directed optimized sociograms to visually assess convergence as the distance between points is a direct and reliable function of their similarity coefficients. The results from this one-mode conversion and MDS plotting are presented in Figure 4. This approach was useful on two levels. First, it demonstrated that, overall, the visual analysis of an optimized graph was a coherent approach. Second, it constitutes a powerful and straightforward triangulation technique for researchers who might wish to try the approach developed here. However, the resulting MDS graphs are not as rich as two-mode graphs produced through force-directed optimization. Information about the degree (salience) of schemata is lost, as is the connection between informants and schemata. For this reason, while one-mode similarity coefficients MDS graphs offer an interesting approach, we believe two-mode graphs produced through force-directed optimization are more promising.

Representing node-based characteristics. To facilitate visual analysis, we plotted the graphs such that node shape corresponded to the informant's role in the healthcare system (physicians as squares, administrators as triangles, nurses as circles, and pharmacists as lozenges), node size corresponded to degree (number of ties connected to one node), and node colour to the node type and the four topics into which all schemata were divided (informants in grey, strengths in purple, problems in red, solutions in green, and influential groups in yellow).

For our study, we produced a set of such sociograms, some including all types of nodes and others displaying only informant nodes and schema-based nodes pertaining to a single topic (i.e., sociograms for informants-to-strengths, informants-to-problems, informants-to-solutions, and informants-to-influential groups). Likewise, we drew the same set of

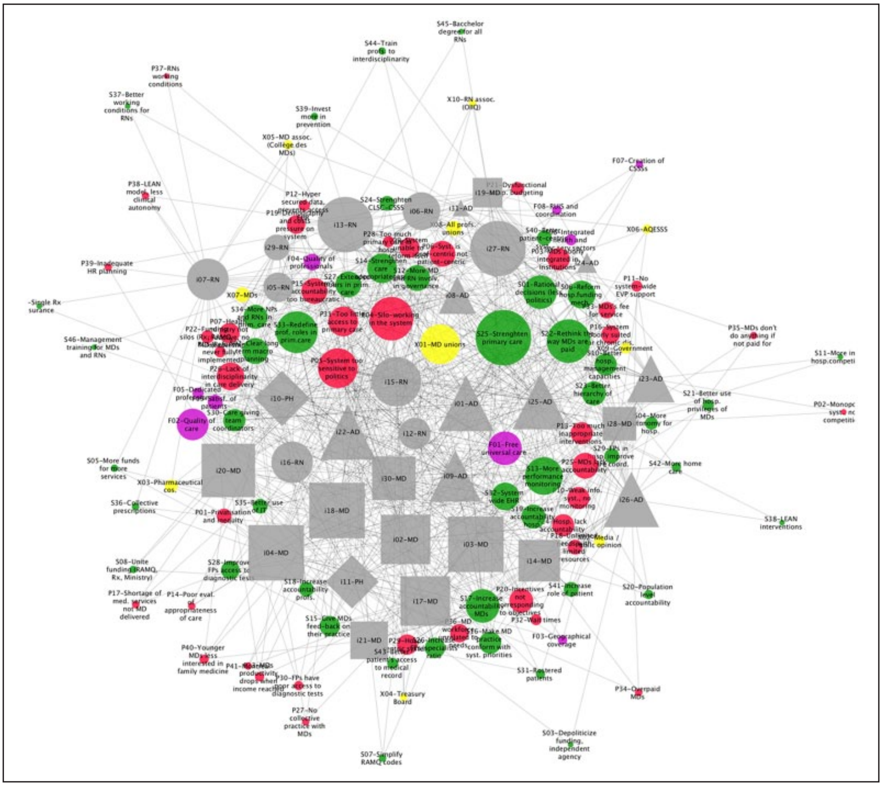

Figure 1. All nodes.

topic-based sociograms for each of the four informant categories (physicians-to-problems, physicians-to-solutions, administrators-to-strengths, etc.). Those sociograms were exported in a scalable graphic format and were the basis for the visual analysis.

Analysis

For the sake of brevity, we do not present here the detailed analysis of the results from this research project nor all the sociograms created. We offer highlights of our findings to demonstrate what our approach produces. The first sociogram (Figure 1) represents the whole dataset.

Given that our dataset included 106 schemata-based nodes and 31 informant-based nodes, the theoretical maximum degree for informant-based nodes was 106, while the maximum degree for any given schema was 31. In their interviews, informants discussed on average 28 different schemata, ranging from 12 (11% of total) to 43 (41%). Schemata occurrence ranged from two (6% of informants) to 28 (90% of informants raised one given solution). Because we have more nodes representing schemata than informants,

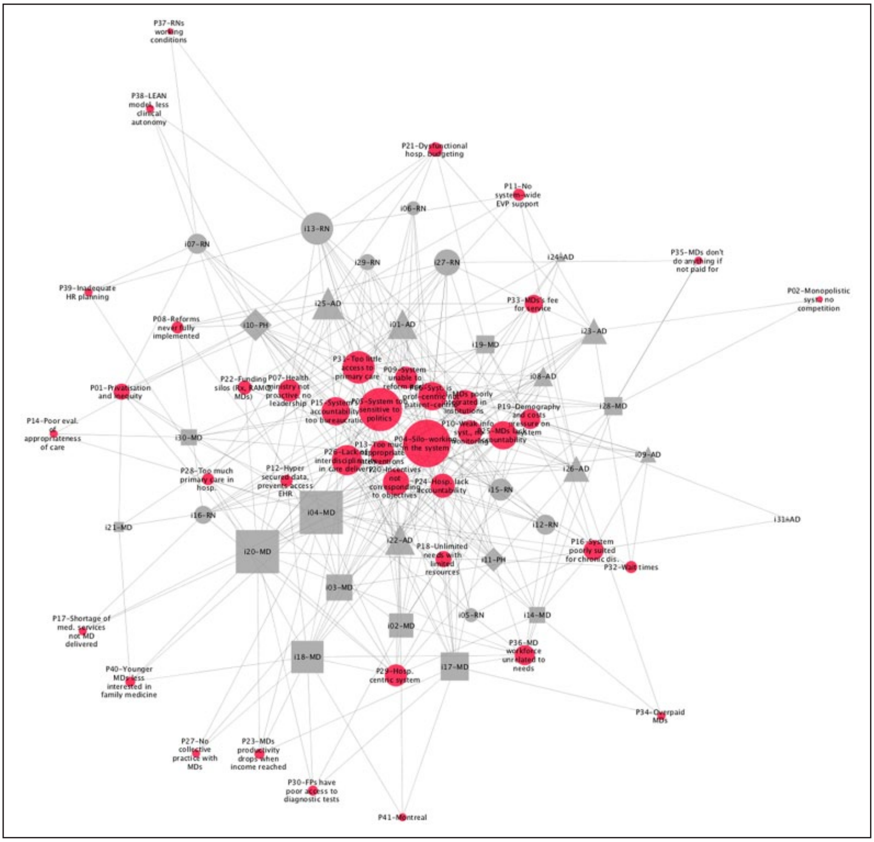

Figure 2. Problems of the healthcare system.

the latter constitute the core of the all-nodes graph, and their relative size compared to schemata-based nodes is large.

One observation regarding this first graph is that all nurses (grey circles) are on the top-left side of the sociogram, while all administrators (grey triangles) are on the bottom-right. This indicates that opinions about the healthcare system go beyond individual variations and are influenced by group structural differences. Conversely, although physicians are mostly on the bottom, they are not as strongly clustered in any one part of the graph, which indicates greater variation in the opinions expressed by this category of informant. MDS of one-mode similarity coefficients gives comparable clustering results. The distribution of informants also means that the opinions (schemata) expressed by nurses will, on average, tend to appear on the top-left side of the graph while those of administrators will appear in the bottom-right. For example, most informants identified physicians as being the most influential group in health-related policy-making. Nevertheless, nurses tended to state this opinion by pointing fingers at physicians in general (node X07, occurrence of 7) or physicians' associations (X05,

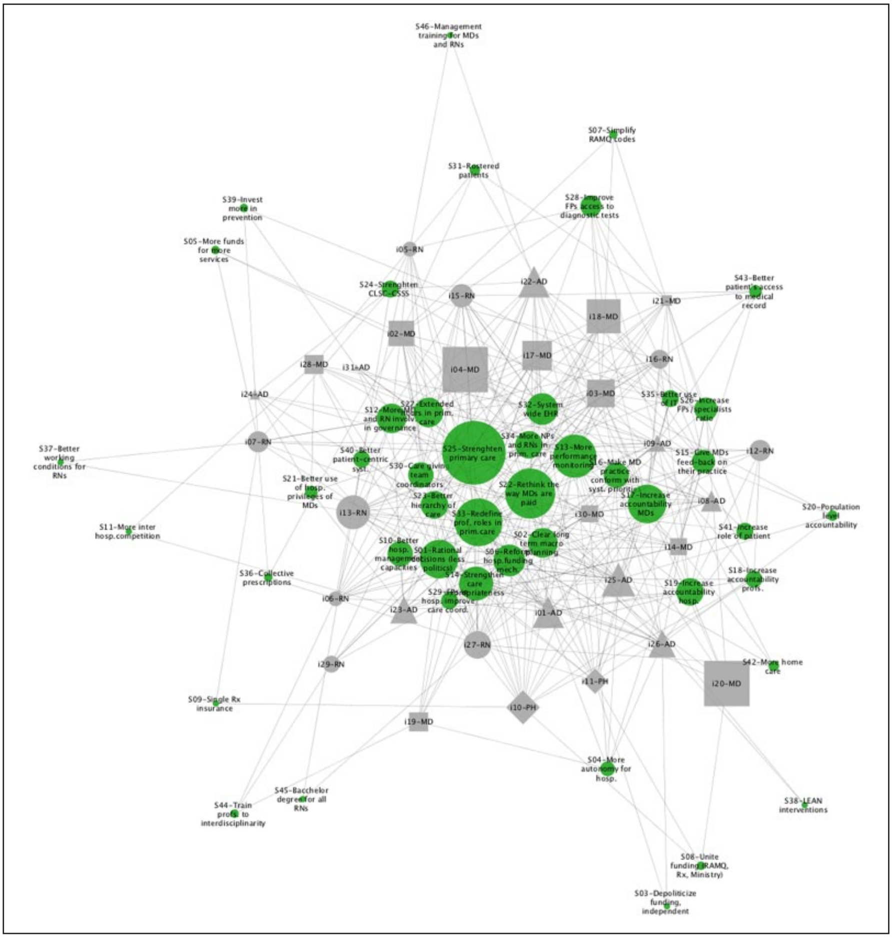

Figure 3. Solutions to reform the healthcare system (two-mode force-directed graph).

occurrence of 4), whereas informants in other categories named specifically physicians' unions (X01, occurrence of 20). This trend explains why the yellow nodes X05 and X07 are located on the top-left side of the graph, closer to the nurses, while the much bigger yellow node X01 is almost central. Such group-based clustering is discussed in more detail with regard to Figures 2, 3 and 4 below.

A second interesting graph is limited to the schemata dealing with healthcare system problems. As we discussed earlier, given the characteristics of our data, informant nodes were positioned at the center of the all-nodes sociogram. However, when the focus was on problems, the balance of the sociogram shifted. On average, informants discussed 10 different problem-related schemata (min: 0 and max: 20, a range of scores denoting informants with extreme tendencies toward optimism or cynicism). Conversely, the most salient problem was raised by 22 informants and the least salient, by two (with an average of seven to eight).

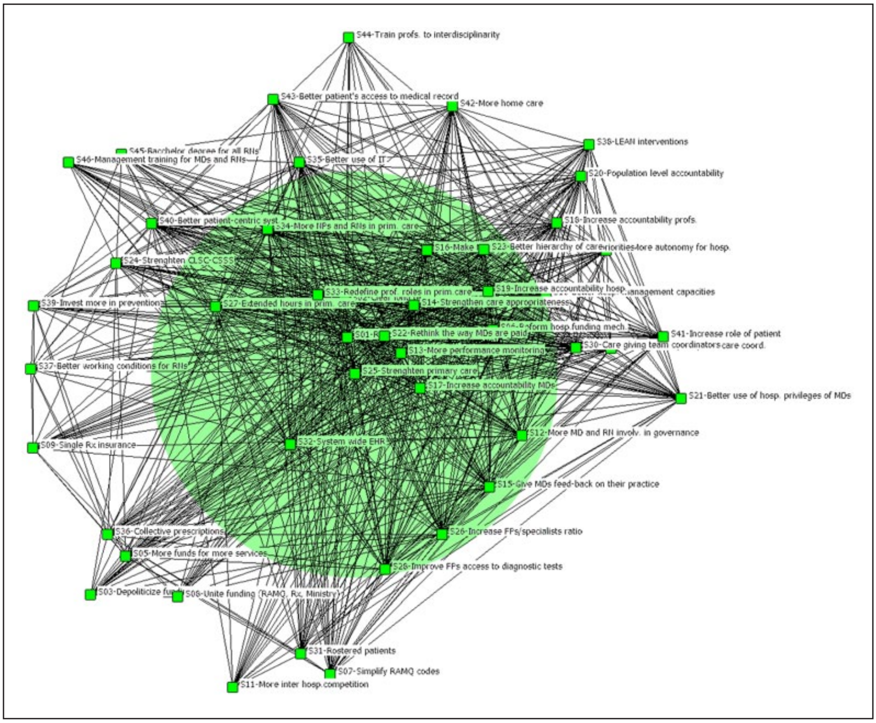

Figure 4. Solutions to reform the healthcare system (one-mode similarity coefficients MDS graph).

This resulted in a core cluster of problem-related schemata at the center of the sociogram. This graph demonstrated that P33 (MDs' fee-for service remuneration), located in the top-right corner, is a relatively salient problem but does not have consensus among informants. It also showed that P04 (working in silos in the health system) and P05 (the healthcare system being too sensitive to politics), the largest nodes located at the center of the graph, were both the most salient and most consensual problems identified by our informants. In examining the problem-related sociogram, we also noticed there was no major clustering of any specific category of informants around any one specific problem.

The solution-related sociogram shared many of the characteristics of the problem-related sociogram but proved to be the most interesting one for our project. First, some nodes had a remarkable level of salience. The most salient, raised by 90% of the informants, had to do with the importance of strengthening the system's primary care capacity (S25). This idea's relative importance and the consensus around it were already obvious as we were conducting interviews. What was not obvious, however, was that there would be a tight cluster of nodes at the center of the graph which, taken together, would constitute a coherent policy proposal. As distances can be directly interpreted on MDS graphs, we added on Figure 4 a circle from node S25 to show this cluster of primary-care interdependent nodes. Both graphs show, in fact, that all the solution-nodes clustered in the center were linked to the overarching idea of strengthening primary care, such as redefining professional roles in primary care (S33), implementing a unified system-wide electronic health record system (S32), rethinking the way physicians are paid (S22), relying more on nurses and nurse practitioners in primary care (S34), integrating care with team coordinators (S30), and hierarchizing care (S23).

With the visualization technique used here, two non-trivial observations could be made. First, it was possible to highlight and discriminate between policy options that mustered a high level of consensus, and were thus positioned at the center of the graph, and the less consensual ones pushed toward the periphery. Second, we were able to identify a surprisingly coherent policy proposal emerging from the interdependent and consensual solutions clustered together at the center of the graph. Once the data are optimized in a visual form, these observations may seem obvious. However, the authors can vouch for the fact that, while working on the data in a narrative form, neither of these two observations was trivial, even for seasoned researchers who conducted many of the interviews and read every transcript several times over.

The visualization also allowed us to assess the plausibility of some of the underlying hypotheses of the research project. For example, we posited that stakeholders holding similar official positions in the healthcare system would tend to present similar opinions. The optimized sociograms revealed some instances where this was the case. However, overall we were surprised by the level of convergence among the different stakeholders' opinions regarding the healthcare system's strengths, problems, and solutions. This example demonstrates how we were able to use our analytical method to enable groupings to emerge empirically rather than drawing conclusions ourselves based on *a priori* group definitions (Marin and Wellman, 2011: 14). We were thereby able to assess empirically whether views expressed by the categories of doctors, nurses, and administrators converged or differed.

Discussion

To summarize our approach: We took narrative data and used a schemata-based coding technique to transform them into a relational form. This process allowed us to use SNA software as well as optimization algorithms derived from mathematical graph theory. The resulting sociograms then constituted our basis for visual analysis and interpretation. In other words, quantitative methods (graph theory) were used to optimize the presentation of narrative data to facilitate their qualitative (visual) interpretation.

This summary of our approach is, in itself, a strong argument to support our point that reliance on quantitative, qualitative, or mixed methods is very much independent of the nature of the data at hand. We argue that the SNA-based method presented here opens up a variety of analysis options and is therefore an interesting addition to other common approaches for making sense of narrative data. It is important to stress that the visualization techniques we propose are complementary to qualitative discourse analysis. The graphs enhanced our comprehension of the data only because we already had an in-depth understanding of them gained from previous projects, through conducting interviews and coding transcripts. Looking at optimized graphs without a good prior understanding of the narrative data would likely bring little to the analysis. As such, we are not arguing here that the visual analysis of graphs should replace qualitative analysis of narrative data, but rather that such graphs provide a powerful complement and support to qualitative analysis (Healy and Moody, 2014).

This method also carries its share of challenges. First, although we believe it is useful for testing the plausibility of some broad hypotheses, such as the convergence or divergence of informants' views, no meaning should be attributed to the individual position of nodes. Individual nodes' position will vary depending on the optimization algorithms chosen. Moreover, the plausibility of any hypothesis should be assessed by comparing qualitative interpretations of the original narrative data against information derived from the graphs. At this level, using both optimized two-mode graphs and MDS of one-mode similarity coefficients can also improve validity. Second, the data collection process presented here and in most research using SNA is not materially different from typical qualitative data collection methods and consequently does not preclude the limitations and biases that qualitative research entails. For example, whether the mapping of relations will be accurate or whether missing links/relations will influence interpretation is contingent on the data collection tools chosen by the researcher. Third, codes are based on a researcher's interpretations and are open to the same methodological questions that arise in other forms of narrative content analysis (for example, the codes might be too large to describe something relevant or too specific to be able to draw conclusions). As with any inductive research process, this method requires continuous critical reflexivity. It should also be thought of as a flexible method to be adapted to each specific dataset and research topic.

The transformation of narrative data into relational data involved in this method is also not without compromises. Using schemata representing the complex thoughts and ideas expressed by the informants as the main unit of analysis rather than revisiting the unabridged interview transcripts does imply a certain information loss. Nevertheless, doing so focuses attention more directly on the relations between ideas and persons rather than on 'individual embodied persons' (Mohr and White, 2008: 489). It is based on the idea that the narrative discourse produced by the informants not only has *sui generis* meaning but also makes sense in relation to other discourses. In this sense, we believe that ideas about healthcare and actors' positions in the healthcare structure form a whole system that should be examined from a holistic perspective (Borgatti and Everett, 1997). Our method allows us to look at the healthcare system as 'organized around cultural networks, relational structures that link meanings, values, stories, and rhetorics together into various structured configurations' (Mohr and White, 2008: 489). For example, by displaying the relations between schemata and informants in a sociogram, we realized that, when taken together, the schemata corresponding to solutions to improve the healthcare system actually constituted a coherent model of healthcare organization (see larger central nodes in Figure 3 and 4). Groupings that were not self-evident among actors occupying different positions in the healthcare structure also emerged empirically out of the visual representation of our research (Marin and Wellman, 2011).

Given time, skill, and energy, it might be possible to arrive at the same conclusions without the mathematical tools available in SNA. However, our case example demonstrates that visual representation and optimization algorithms greatly facilitate interpretation, even with mid-scale datasets such as ours.

There has been a trend in SNA scholarship to single out this method from other analytical tools and to focus on developing theories of social network properties (for example, reproducible patterns of social network). This perspective applies mostly to specifically relational datasets, and the SNA scholarship in this area is already considerable (Mohr and White, 2008). However, methods such as the one proposed here that examine both structures of actors and structures of meanings and ideas can be applied to a much wider variety of datasets. Moreover, in our study the narrative data was collected from interviews, but other textual material could very well undergo the same analytical process (e.g. using narrative data from articles to form a network structure comprising positions about a given topic and from different newspapers). The method could hypothetically also be applied to a single source of narrative data while applying a chronology (e.g. collected works of a scholar used to create a network of themes covered in different time periods).

Conclusion

As discussed earlier, we believe the method proposed here can be productively combined with other quantitative and qualitative methods to enrich the analysis of narrative data. Given the recent development of powerful, user-friendly, and open-source SNA software, anyone working with narrative data can apply the method quite straightforwardly.

Software-based coding has proven to be extremely powerful for organizing large datasets and producing coding statistics, occurrence counts, etc. However, software-based coding of narrative data remains surprisingly unhelpful for making sense of data inductively. In our experience, even when coding software is used, sense-making emerges idiosyncratically from human processing of the data in its narrative form through repeated readings of the text and, hopefully, from interconnection with an underlying conceptual framework. We believe the method presented here can be useful to inform and support this inductive sense-making. It is not intended to replace the human processing of narrative data, but rather to support it.

In conclusion, we would like to point out that, since the method proposed here is based on coding techniques that are very similar to what is now the mainstream approach used in popular qualitative data analysis software, our method could quite easily be integrated directly into those narrative data analysis software programs. This would give users the possibility to obtain optimized sociograms instantly without additional efforts. In our view, the potential for applying a method derived from social network analysis to work with narrative data, still largely unexplored, offers a relevant and fertile area for future study.

Funding

The authors want to acknowledge the financial support of the Canadian Institutes for Health Research (CIHR) for the lead author's Research Chair as well as for the operating grant (no. 272944).

Note

1. We are grateful to the anonymous reviewer of this article who suggested this very useful additional validation approach.

References

Barnes J (1954) Class and committees in a Norwegian Island parish. *Human Relations* 7(1): 39–58.

Bearman PS and Stovel K (2000) Becoming a Nazi: a model for narrative networks. *Poetics* 27(2): 69–90.

Bellotti E (2014) *Qualitative Networks: Mixed Methods in Sociological Research*. Abingdon: Taylor & Francis.

Bidart C and Lavenu D (2005) Evolutions of personal networks and life events. *Social Networks* 27(4): 359–376.

Blommaert J and Bulcaen C (2000) Critical discourse analysis. *Annual Review of Anthropology* 29: 447–466.

Borgatti SP (2009a) Things I have wanted to say about the analysis of 2-mode networks but hadn't had the opportunity to. Until now. Conference and Workshop on Two-Mode Social Network Analysis. October 1st. VU University: Amsterdam.

Borgatti SP (2009b) Two-mode concepts in social network analysis. In: Meyers RA (ed) *Encyclopedia of Complexity and System Science*. New York: Springer.

Borgatti SP and Everett MG (1997) Network analysis of 2-mode data. *Social Networks* 19: 243–269.

Borgatti SP, Everett MG and Johnson JC (2013) *Analyzing Social Networks*. London: SAGE.

Bott E (1955) Urban families: conjugal roles and social networks. *Human Relations* 8(4): 345–384.

Bourdieu P (1982) *Ce que parler veut dire: l'économie des échanges linguistiques*. Paris: Fayard.

Chouliaraki L and Fairclough N (1999) *Discourse in Late Modernity*. Edinburgh: Edinburgh University Press.

Cline MS, Smoot M, Cerami E, et al. (2007) Integration of biological networks and gene expression data using Cytoscape. *Nature Protocols* 2(10): 2366–2382.

Contandriopoulos D, Abelson J, Lamarche P, et al. (2012) The visible politics of the privatization debate in Quebec. *Healthcare Policy* 8(1): 67–79.

Crossley N (2010) The social world of the network. Combining qualitative and quantitative elements in social network analysis. *Sociologica* 4(1): 1–34.

DiMaggio P (1997) Culture and cognition. *Annual Review of Sociology* 23(1): 263–287.

Domínguez S and Hollstein B (2014) *Mixed Methods Social Networks Research: Design and Applications*. Cambridge: Cambridge University Press.

Edwards G (2010) *Mixed-method approaches to social network analysis*. Discussion Paper. NCRM.

Fairclough N (2013) *Critical Discourse Analysis: The Critical Study of Language*. New York: Routledge.

Franzosi R (1994) From words to numbers: a set theory framework for the collection, organization, and analysis of narrative data. *Sociological Methodology* 24: 105–136.

Freeman L (2004) *The Development of Social Network Analysis*. Vancouver, BC: Empirical Press.

Fruchterman TMJ and Reingold EM (1991) Graph drawing by force-directed placement. *Software: Practice and Experience* 21(11): 1129–1164.

Gastner MT and Newman MEJ (2006) The spatial structure of networks. *The European Physical Journal B – Condensed Matter and Complex Systems* 49(2): 247–252.

Hardy C, Palmer I and Phillips N (2000) Discourse as a strategic resource. *Human Relations* 53: 1227–1248.

Healy K and Moody J (2014) Data visualization in sociology. *Annual Review of Sociology* 40(1): 105–128.

Heath S, Fuller A and Johnston B (2009) Chasing shadows: defining network boundaries in qualitative social network analysis. *Qualitative Research* 9(5): 645–661.

Hollstein B (2011) Qualitative approaches. In: Scott J and Carrington PJ (eds) *The SAGE Handbook of Social Network Analysis*. London: SAGE, 404–417.

Hsieh H-F and Shannon SE (2005) Three approaches to qualitative content analysis. *Qualitative Health Research* 15(9): 1277–1288.

Kadushin C (2012) *Understanding Social Networks: Theories, Concepts, and Findings*. Oxford: Oxford University Press.

Keim S, Klärner A and Bernardi L (2009) Qualifying social influence on fertility intentions: composition, structure and meaning of fertility-relevant social networks in western Germany. *Current Sociology* 57(6): 888–907.

Knoke D and Yang S (2008) *Social Network Analysis*. Thousand Oaks, CA: SAGE.

Knox H, Savage M and Harvey P (2006) Social networks and the study of relations: networks as method, metaphor and form. *Economy and Society* 35(1): 113–140.

Krippendorff K (2012) *Content Analysis: An Introduction to its Methodology*. Thousand Oaks, CA: SAGE.

Light R (2014) From words to networks and back: digital text, computational social science, and the case of presidential inaugural addresses. *Social Currents* 1(2): 111–129.

Lorrain F and White HC (1971) Structural equivalence of individuals in social networks. *The Journal of Mathematical Sociology* 1(1): 49–80.

Marin A and Wellman B (2011) Social network analysis: an introduction. In: Scott J and Carrington PJ (eds) *The SAGE Handbook of Social Network Analysis*. London: SAGE, 11–25.

Mauws MK and Phillips N (1995) Crossroads understanding language games. *Organization Science* 6(3): 322–334.

Mckether WL, Gluesing JC and Riopelle K (2009) From interviews to social network analysis: an approach for revealing social networks embedded in narrative data. *Field Methods* 21(2): 154–180.

Mische A (2003) Cross-talk in movements: reconceiving the culture-network link. In: Diani M and McAdam D (eds) *Social Movements and Networks: Relational Approaches to Collective Action*. Oxford: Oxford University Press, 258–280.

Mitchell JC (1969) Social Networks in Urban Situations: Analyses of Personal Relationships in Central African Towns. Institute for Social Research, University of Zambia/University of Manchester Press: Manchester.

Mohr JW and White HC (2008) How to model an institution. *Theory and Society* 37(5): 485–512.

Morgan DL (1993) Qualitative content analysis: a guide to paths not taken. *Qualitative Health Research* 3(1): 112–121.

Newman M, Barabasi A-L and Watts DJ (2006) *The Structure and Dynamics of Networks*. Princeton, NJ: Princeton University Press.

Palla G, Derényi I, Farkas I and Vicsek T (2005) Uncovering the overlapping community structure of complex networks in nature and society. *Nature* 435(7043): 814–818.

Prell C (2012) *Social Network Analysis. History, Theory & Methodology*. London: SAGE.

Radcliffe-Brown AR (1940) On social structure. *The Journal of the Royal Anthropological Institute of Great Britain and Ireland* 70(1): 1–12.

Riff D, Lacy S and Fico F (2014) *Analyzing Media Messages: Using Quantitative Content Analysis in Research* (third edition). Abingdon: Taylor & Francis.

Saussure F (1955) *Cours de Linguistique Générale*. Paris: Payot.

Scott J (2000) *Social Network Analysis: A Handbook*. London: SAGE.

Scott J and Carrington PJ (2011) *The SAGE Handbook of Social Network Analysis*. London: SAGE.

Shannon P, Markiel A, Ozier O, et al. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research* 13(11): 2498–2504.

Small ML (2011) How to conduct a mixed methods study: recent trends in a rapidly growing literature. *Annual Review of Sociology* 37(1): 57–86.
Smith T (2007) Narrative boundaries and the dynamics of ethnic conflict and conciliation. *Poetics* 35(1): 22–46.
Tausczik YR and Pennebaker JW (2010) The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology* 29(1): 24–54.
Van Dijk TA (2011) *Discourse studies: A Multidisciplinary Introduction*. London: SAGE.
Wasserman S and Faust K (1994) *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.
Wodak R and Meyer M (2009) *Methods for Critical Discourse Analysis*. London: SAGE.

Author biographies

Damien Contandriopoulos is a Professor at the Faculty of Nursing, University of Montreal, and Researcher at the University of Montreal Public Health Research Institute. He holds a Canadian Institutes of Health Research (CIHR) Applied Public Health Chair on Health Policy and Evidence.

Catherine Larouche is a PhD Candidate in Anthropology, McGill University, and Research Assistant at the University of Montreal Public Health Research Institute.

Mylaine Breton is an Assistant Professor in the Department of Community Health Sciences, University of Sherbrooke, and Researcher at the Charles-LeMoyne Hospital Research Center, University of Sherbrooke.

Astrid Brousselle is a Professor in the Department of Community Health Sciences, University of Sherbrooke, and Researcher at the Charles-LeMoyne Hospital Research Center, University of Sherbrooke.