Journal of Social Structure

JoSS Article: Volume 14

Using Visualizations to Explore Network Dynamics

Kar-Hai Chu, karhai.chu@usc.edu
Heather Wipfli, hwipfli@med.usc.edu
Thomas W. Valente, tvalente@usc.edu

Institute for Prevention Research, Department of Preventive Medicine, Keck School of Medicine, University of Southern California

Abstract

Network analysis has become a popular tool to examine data from online social networks to politics to ecological systems. As more computing power has become available, new technology-driven methods and tools are being developed that can support larger and richer network data, including dynamic network analysis. This timely merger of abundant data and cutting edge techniques affords researchers the ability to better understand networks over time, accurately show how they evolve, find patterns of growth, or study models such as the diffusion of innovation. We combine traditional methods in social network analysis with new innovative visualizations and methods in dynamic network studies to explore an online tobacco-control community called GLOBALink, using almost twenty years of longitudinal data. We describe the methods used for the study, and perform an exploratory network study that links empirical results to real-world events.

Keywords

Social network analysis, dynamic visualization, longitudinal analysis

Acknowledgements

This study was supported by NIH grant number 1R01CA157577 (PI: Thomas W. Valente) and training grant T32CA009492 from the National Cancer Institute/National Institutes of Health.
An earlier version was presented at the XXXIII annual INSNA meeting, Hamburg, Germany.
We would like to thank the International Union for Cancer Control for its assistance with this study.

1. Introduction

Visualization tools have been a strong component of scientific progress in various fields. In many cases, data summarized as graphs or charts can help clearly represent ideas. In other examples, concepts have become associated with a particular image that originated from research, from the miniscule double helix twisted ladder of DNA to the large spiral arms of the Milky Way. A great deal of information can be derived from simple images, whether viewing a line graph of company stock or a pyramid view of a food chain.

Network graphs, in particular, are a useful tool that can help model relations, summarize data, and represent abstract concepts in a clear and intuitive way. The value of using network graphs to visualize data has been applied in different fields, and has helped improve our knowledge of disease spread (Christakis and Fowler 2010), international telecommunications (Barnett 2001), ecological systems (Stefano, Alonso and Pascual 2008), social networks (Moody and White 2003), health studies (Valente 2010), among many others. Social network analysis (SNA) often uses a *sociogram* to clarify different concepts. Sociograms are network graphs in which nodes represent actors and ties represent relationships between them.

The sociogram is a powerful analysis tool, helping researchers identify points of interest such as clusters (Newman and Girvan 2004), boundary spanners (Levina and Vaast 2005), central and peripheral layers (Borgatti and Everett 2000), and other structural properties that otherwise would not be obvious in numeric data (e.g. an adjacency matrix). Today, there are online communities that form around every conceivable topic, so it is no surprise that SNA has become popular for online social network research.

Growing in parallel with SNA is the availability of different software tools. Since Moreno's (1932) small hand drawn examples, modern computer technology can now create networks with 10's of millions of users (Mislove, Massimiliano, Gummadi, Drushel and Bhattacharjee 2007). The development of SNA software has aided SNA research, as increased computing power has enabled fast complex calculations and supported large-scale network analyses (e.g. visualizing million node networks). Researchers can conduct studies based on network structures, and many of the calculations and measurements are made immediately available. Methodological developments are often paired alongside certain software, such as exploratory analysis using Pajek (de Nooy, Mrvar and Batagelj 2005). Other software packages each have their own benefits, such as UCINet's[1] easy support of many SNA tools, or the statnet package built into the freely available R environment[2], offering great flexibility and statistical analyses.

Given the power of SNA, there are still gaps that have only recently started to be addressed. For example, sociograms are, by nature, static representations. They are snapshots of a network in a single moment in time, giving no hints as to how or why the network developed into a particular structure, or what it could potentially become. More studies into the evolution of social networks would be beneficial for research, especially in online communities, which can grow at tremendous speeds.

This paper applies SNA and dynamic network visualizations to study the growth and evolution of GLOBALink, an online network focused on global tobacco control. In analyzing GLOBALink data collected over a 20-year period, we are not only able to visualize the membership network over time, but can also link shifts in the network to major political, social, and economic changes that occurred in the global tobacco control community. These events include major cultural shifts regarding tobacco use in high-income western countries; the negotiation of the first public health treaty, the Framework Convention on Tobacco Control; and major philanthropic donations to combat tobacco use in low- and middle-income countries. The paper begins with an overview of the methods, data preparation and software tools used, followed with numerical and visual results, including a movie representing the evolution of the community. Finally, we draw conclusions related to the network trends apparent through the software visualizations, major external influences on the development of the network, and the implications these initial results have for future network analyses focused on evaluating the role and impact of social and political influences on network formation and evolution.

2. Background

2.1. SNA visualizations

Within social network analysis, researchers have recognized the value in emphasizing important features of social structures, the similarities and differences in positions occupied by the actors, searching for groups and positions, and understanding the patterns that link sets of actors (Freeman 2000). Freeman noted the strength of the sociogram

---

[1] https://sites.google.com/site/ucinetsoftware/home
[2] http://www.r-project.org/

as a method of exploration, and also predicted that as computing processing power and storage continued increasing, there would be a growth in graph-generating software. While browser-based Java applets and VRML tools did not become as popular as he predicted, there are many standalone network analysis software packages that have been developed.

Over a decade later, we are still learning how to visualize social networks. Recently, Correa and Ma (2011) identified 4 types of social network visualization: structural, semantic, temporal, and statistical. They also describe how layouts should satisfy the needs of readability, clusterability, and trustworthiness. Like Freeman, their research and techniques build on a process of adopting new technologies to advance network visualizations. The integration of modern computing hardware has helped network analysts examine extremely large datasets. In parallel, new software are being developed that not only provide visualizations, but also support more research-oriented practices. Packages such as SocialAction have been developed that better integrate classical methods in exploratory data analysis and statistics with SNA visualizations (Perer and Shneiderman 2008). SocialAction was used to find different levels of partisanship in US senators by interactively filtering the data on various statistical measures. Along with other visualization-focused technologies (e.g. Gephi, ORA, NetLogo), new tools are being developed that enable network graphs to be integrated into different types of research.

2.2. *Dynamic Network Analysis (DNA)*

Dynamic network analysis can provide an aid to longitudinal SNA research. However, as a relatively young field, many aspects have not been explored and there are few standards that have been established. One path taken was to treat network edges as probabilistic, and use multi-agent systems to study network evolution (Carley 2003). Carley redefined the traditional sociogram by adding probabilistic parameters on the edges, providing a quantification of the likelihood they will form. Individual nodes were also given more emphasis; they are treated as agents, and can potentially impact how a network will develop. Terrorist networks are a good example where prediction is vital (Krebs 2002), and one in which link prediction has important applications (Carley, Dombroski, Tsvetovat, Reminga and Kamneve 2003; Liben-Nowell and Kleinberg 2006). Another visualization method is to explicitly use the time and order of social interactions to build the network (Berger-Wolf and Saia 2006).

Researchers have also applied DNA methods to study other forms of longitudinal networks. Kossinetts and Watts (2006) examined the stability of bridges, defined as connections outside one's circle of acquaintances, and measured how social ties were created and dissolved over time. Barabasi and colleagues (2002) applied DNA methods to find unique properties in an evolving citation network that differed from classic models. Indeed, various forms of citation and co-authorship networks offer a wealth of reliable data with which to study the evolution of networks. New layouts and metric computations are constantly being developed from these dynamic network studies (Brandes and Pich 2012).

2.3. Dynamic network visualization

The study of dynamic networks greatly benefits from visualizations that can illustrate ideas and concepts not immediately visible in a static sociogram. In fact, "The ability to see data clearly creates a capacity for building intuition that is unsurpassed by summary statistics" (Moody, McFarland and Bender-deMoll 2005). Moody and others' research emphasizes how the ability to see data can be superior to summary statistics, and illustrates the need to visualize how networks develop and change over time. Additionally, they lay the foundation of how dynamic network visualizations should be presented (e.g. differentiating between discrete and continuous time), and recommend visualization and analysis be interactive. These theoretical ideas were developed in parallel with SoNIA, a software package for visualizing dynamic network data (Bender-deMoll and McFarland 2006). Other researchers have continued to focus on studying different properties of dynamic networks, e.g. the evolution of subgroups (Falkowski, Bartelheimer and Spiliopoulou 2006), effects of network topology and organizational structure over time (Kossinets and Watts 2006), detecting and predicting statistically significant changes in a network over time (McCulloh and Carley 2011), and new visualization methods using shortest-path computations (Brandes and Pich 2012).

2.4. Measuring online communities

Studies of online communities have expanded as the number of online communities continues to increase. In particular, social networking sites have become a valuable source of data for different types of studies. User actions can be easily and accurately collected over extended periods of time, for large numbers of websites. Each site can have distinct characteristics, distinguished by the purpose for which people joined, the type of interactions that occur, or the media available for use by members (Chu and Suthers 2013). These variations afford many opportunities to study the similarities and differences between online communities (Hether, Murphy and Valente in-press).

There are different metrics by which online communities can be compared. However, given how website developers will naturally build unique features to attract users to their particular community, it has become increasingly difficult to compare such metrics across different online communities, e.g. types of friendships in Friendster and MySpace (boyd 2006), social capital in Facebook (Ellison, Steinfield and Lampe 2007), or diffusion in Twitter (boyd, Golder and Lotan 2010; Suh, Hong, Pirolli and Chi 2010) and YouTube (Susarla, Oh and Tan 2012), to name a few.

One option where this is possible is to examine membership patterns for a given online community, e.g. growth rate. Here we can apply more generalized social networking measures to study network structure, use centralization measures which can be viewed independent of the characteristics of the online features in each website. Additionally, we can apply well-established models to these patterns to help study membership patterns in these communities. For example, Backstrom and colleagues (2006) studied membership and growth of weblog site LiveJournal and publication database DBLP by comparing it to a diffusion model (Valente 1995; Rogers 2003; Valente 2005). They examined the members' existing friends, and watched how group growth developed through ties its members had to individuals outside of their groups. Other studies (Firth, Lawrence and Clouse 2006) have similarly compared membership adoption to the diffusion model and found similar results.

SNA metrics and concepts can also be useful in helping to understand actions within online communities. Mislove and others (2007) studied the communities in Flickr, YouTube, LiveJournal, and Orkut, representing some of the largest online communities at the time (over 11.3 million users and 328 million links). They found the communities exhibited a strongly connected core of high-degree nodes, surrounded by many small clusters of low-degree nodes. There was also a high degree of reciprocity in directed user links, leading to a strong correlation between user in-degree and out-degree, with a power law degree distribution. A more recent study by Kairam and colleagues (2012) investigated the Ning network to compare diffusion and non-diffusion membership growth, and found that clustering promotes diffusion growth, although it is more likely to lead to smaller eventual groups.

2.5. Study Sample and Significance

Sociograms continue to be vital for researchers to highlight the important details of a network at a given moment, but expanding the static image to a continuous movie can reveal much more. Just as static graphs can reveal structures and properties that are not visible in numeric data representations, dynamic movies can reveal patterns over time, how a network behaves, and the shift of nodes as new ties are made. These data can answer questions such as how a particular structure came to be formed or destroyed, and possibly predict how it might change in the future.

Online communities can provide large amounts of accurate data, usually based on web server log files. Popular web servers such as Apache[3] provide logs with rich information, including user-identifiers (e.g. IP address), precise timestamps, mouse clicks, text typed into any field, etc. Browser cookies allow servers to remember users across multiple web surfing sessions. These resources offer researchers rich and accurate information in how people communicate and interact in any online medium. It is an ideal source of data for dynamic network studies.

In developing our study, we searched for an online community that had been in place for enough time that we could test not only the SNA and visualization methods and technologies, but would also allow us to examine the impact of social, cultural, and political events on the shape and evolution of the network. Global tobacco control provided an ideal case. GLOBALink, the online global tobacco control community, has been in place for over 20 years. In 1992, the Switzerland-based International Union for Cancer Control (UICC) took over coordination of a small US-based tobacco control network and formed GLOBALink, an online network of tobacco control professionals. Over the following two decades, GLOBALink grew into a large network with members from throughout the world dedicated to controlling tobacco use. GLOBALink’s homepage contains news bulletins, electronic conferences, live interactive chat, and full-text databases (including news, legislation, directories). GLOBALink has been recognized for its ability to unite the global tobacco movement. For example, GLOBALink received the Tobacco or Health Award from WHO for its unique ability to bring together tobacco control advocates, while the tobacco industry recognized the power of GLOBALink in uniting the domestic tobacco control movements, remarking: “But what brings all these groups close together is the excellent communications network that has been built up” (Waller and Lipponen 1997).

GLOBALink’s growth and evolution took place at a time of dramatic change in the access to, and use of, online technology and platforms. Similarly, dramatic social and political shifts also took place in the global tobacco control environment. In fact, global tobacco control has

---

[3] http://httpd.apache.org/

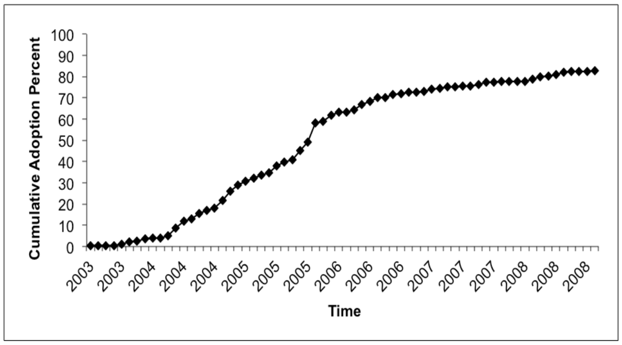

Figure 1: Global diffusion of the Framework Convention on Tobacco Control

been one of the most dynamic areas of global public health in the past two decades. In 1999, the WHO began negotiations on a framework convention on tobacco control, an international treaty aimed at reducing the global burden of tobacco-related death and disease. The 192 member states of the WHO unanimously adopted the Framework Convention on Tobacco Control (FCTC) in 2003 and the FCTC became binding international law less than two years later. Within two years, over 168 states had signed the treaty and 140 states had ratified it (Figure 1). Several existing studies have examined the FCTC ratification process. Wipfli and others (2010) found that the likelihood of FCTC ratification by a country was three times as likely when that country was exposed to other members of ratifying countries via membership in GLOBALink. Since 2006, New York City Mayor Michael Bloomberg and Bill Gates have contributed over $750 million to efforts to implement the FCTC obligations in low- and middle-income countries throughout the world. Consequently, we can analyze both general social networking trends, and tobacco control-specific events, to see which technological or societal events appear to have an impact on GLOBALink membership and relationships.

GLOBALink is also a unique online network in that membership requires a multi-step process by which two existing GLOBALink members are required to vouch for any new applicant. This procedure was originally put in place to prevent tobacco industry employees from participating in GLOBALink. We have not found similar studies that examine a community where 100% of the members require a double internal validation method. As such, we expect that membership patterns will be different than other online communities with open membership. The resulting network should also have a very low likelihood of isolates; as our ties represent advocacy of membership, isolated clusters will not form unless a founding member advocates for a few members who, in turn, do not refer new members. Our study findings consequently can be helpful for those interested in how to study network evolution, as well as for professionals and policy makers interested in how policy decisions may impact the way online communities communicate and work together to address shared problems.

3. Methodology

We performed an exploratory study of GLOBALink members according to their self-reported home country.

3.1. Data

We developed a customized Java application using crawler4j[4], an open source multi-threaded web crawler. The program retrieved the full membership list of all GLOBALink members, and then mined their pages for referral information. Member profiles also contained information for the country that the member represents, which must be chosen from a predefined drop down list. By recursively parsing the referral and country data for each member, we constructed a database of relationships between countries based on membership referrals. However, no personal information is included in this study. The unit of analysis is at the country level, and any member-specific information is aggregated to their listed home country.

We explored new membership patterns in the GLOBALink community to see if they conformed to a diffusion process. We expect that members from different countries will seek membership in GLOBALink in parallel paths of how countries ratify the FCTC. However, because GLOBALink has a unique membership verification process, we do not expect that new membership patterns to be similar to those in online communities as found in other studies (e.g. Backstrom et al. 2006; Firth et al. 2006).

3.2. Visualization Software

There are many popular software tools, such as Pajek and UCINet, which have been used for various types of SNA research. However, they do not provide support for dynamic visualization and are less beneficial for longitudinal analyses. This paper does not attempt to comprehensively compare different software packages that can visualize dynamic networks. There are several available off-the-shelf tools that support some form of longitudinal analysis, including DyNet[5], GUESS[6], SoNIA (Bender-deMoll and McFarland 2006), TeCFlow (Gloor and Zhao 2004), and JUNG (O'Madadhain, Fisher, White and Boey 2003). In this paper, we use Gephi, an open source multi-purpose platform for network visualization.

[4] http://code.google.com/p/crawler4j/
[5] http://www.casos.cs.cmu.edu/projects/DyNet/index.php
[6] http://graphexploration.cond.org/

3.3. *Social network analysis using Gephi*

As described on their website, Gephi is an interactive visualization and exploration platform for all kinds of networks and complex systems, dynamic and hierarchical graphs. Some of its most attractive features are its support of many different native graph formats, real-time interactive features, and easy-to-use interface. Most importantly, it has many supporting features built for dynamic network analysis that incorporate functions such as live filtering, a combination of static and dynamic metrics, a multitude of layouts, and a timeline component that can generate various longitudinal reports. It is an open-source project that is constantly updated with community-contributed plugins. As of this writing, the most current version is beta 0.8.2. More information on the Gephi software can be found in their release paper (Bastian, Heymann and Jacomy 2009) or website[7].

Our web crawler provided information for GLOBALink members that included their home country, the countries of their referring members, and the date when their account was approved. This was stored in a comma-separated values (CSV) file. Custom parsing scripts were run against the CSV file to create two Gephi-readable files: a dynamic nodes list, and a dynamic ties list. The dynamic nodes are a list of all nodes in the network (i.e. countries), with a time interval based on when a member from a country first joins. This was achieved by parsing through all members, grouped by their home countries, and selecting the one with the earliest account approval date. The results were formatted to be imported into Gephi by listing each country, and the time interval in which the node exists. An example of the data follows:

```
“Id” "Time Interval"
"Portugal" "<[1999-05-10,2013-01-01]>"
"Chile" "<[1997-07-20,2013-01-01]>"
```

The dynamic ties are a list of all ties in the network. Unlike with nodes, a tie between two given countries will change over time, i.e. increase in weight as more referrals are made. Thus, there are two time interval components, one that represents the changes in weight, and another that shows the total lifespan. To create this file, we parse the original CSV and create a tie for each referral-referee country pair, and include the date. When a new pair is found, we increment the weight of the tie, and append the new date range. An example of this data:

```
"source" "target" "weight" "Time Interval"
```

[7] https://gephi.org

```
"Australia" "Nepal" <[2001-09-11,2013-01-01,1];> <[2001-09-11,2013-01-01]>

"Switzerland" "Italy" <[1993-01-01,1996-05-13,1];[1996-05-13,1998-10-06,2];[1998-10-06,1999-07-05,3];[1999-07-05,2002-11-19,4];[2002-11-19,2013-01-01,5];> <[1993-01-01,2013-01-01]>
```

These files are imported into a blank Gephi project, through the Data Laboratory. We can now enable the Timeline feature for movie playback. After calculating the dynamic node and edge statistics, we can vary the node size changes in real-time by using the Rank feature to show Dynamic Out-Degree. Similarly, the layout algorithm can be run during the Timeline playback to show how the network structure is changing through the years.

4. Results

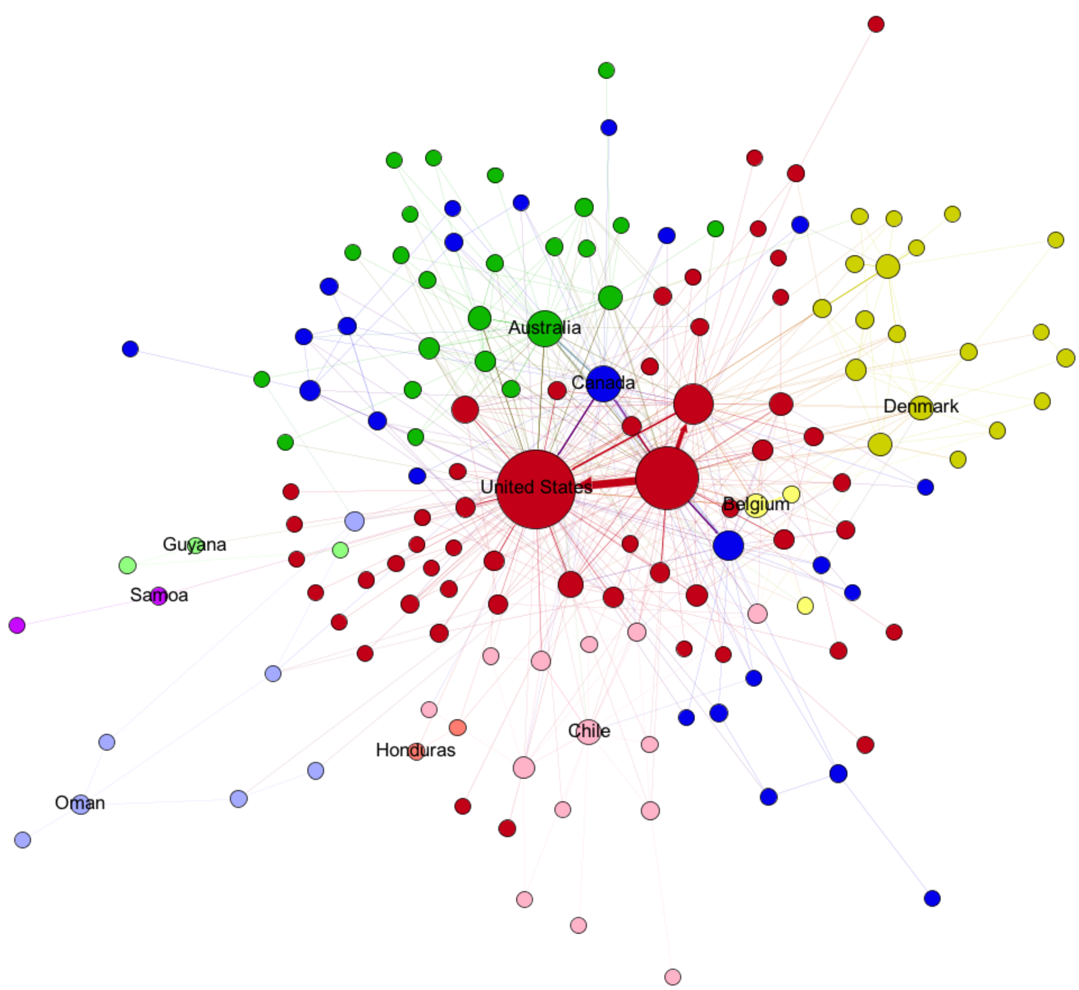

Figure 2: GLOBALink referral network

4.1. Static network

We describe the results of the network analyses using GLOBALink membership referral data. Figure 2 is a final static snapshot of the network as of January 1, 2013.

4.1.1. Description

Nodes represent the countries for all the members in GLOBALink. A directed edge is created when a member from one country (source) vouches for a member from another country (target). Node size represents out-degree, i.e. the total number of target country referrals made from the source country. Node color represents modularity class, a classification of the community a node belongs to (more details in the Discussion section). From each of the ten modularity classes that were detected, the highest out-degree node in each cluster is labeled. The edge size represents the count of referrals between two countries, in either direction. Edge colors are derived from the color of their source nodes.

The layout of the network is created using Gephi's ForceAtlas2 (FA2) algorithm, which created an easy-to-interpret graph. The FA2 algorithm is continuous and optimized for speed (suitable for dynamic graphs, as it will efficiently update in real time) and offers various options to help fine-tune the results. More details can be found on Gephi's website.

4.1.2. Statistics

There are 152 nodes connected by 611 edges. Since we are focusing on a diffusion model and a longitudinal view of how the network evolves, self-loops – edges that are created by members vouching for new members from their home countries – are ignored. This reduces our edge count to 519, approximately 85% of the complete network. Table 1 reports other network metrics. All measurements are computed with directed edges, and with self-loops removed.

Table 1: General statistics of the GLOBALink referral network

Average degree	3.414
Average weighted degree	6.909
Graph density	0.027
Modularity Q	0.305 (10 communities)
Clustering coefficient	0.310
Average path length	2.671

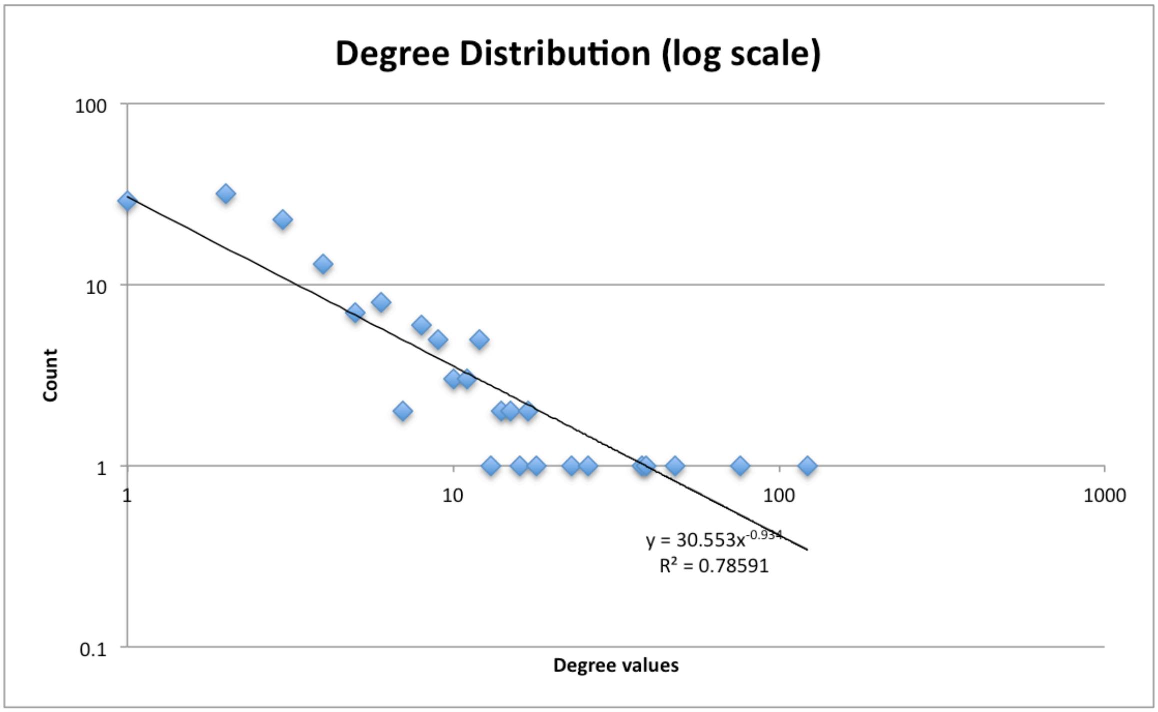

Figure 3: Degree distribution of nodes, in a log-log graph. Exponent of power function is 0.934

Figure 3 is the degree distribution of the network, and follows a standard power law.

4.2. Dynamic network

4.2.1. Movie

By enabling Gephi's Timeline feature, we created a movie[8] to observe the network's evolution over time, in the 20-year period from January 1, 1993 to January 1, 2013. The video runs approximately 31 seconds. In several instances, some nodes will float out of the screen, but they are all returned as the movie progresses. Node color and size continue to represent modularity class and out-degree, respectively.

The video can be viewed here

[8] The video was captured using Screencast-o-matic, found at http://www.screencast-o-matic.com

4.2.2. Graphs

Several graphs were generated based on the Timeline data from the dynamic network. Figure 4 and Figure 5 depict the total number of nodes and edges over the 20-year period. Figure 6 represents the global (average) clustering coefficient over the same time.

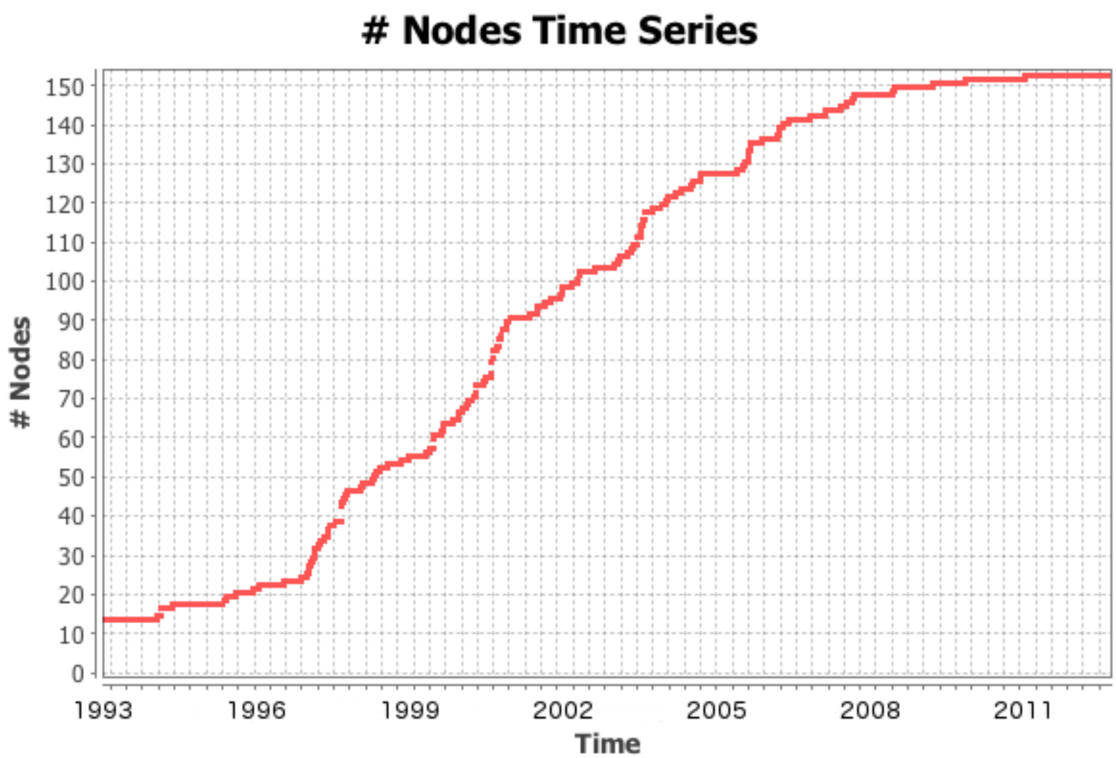

Figure 4: Total number of nodes over time

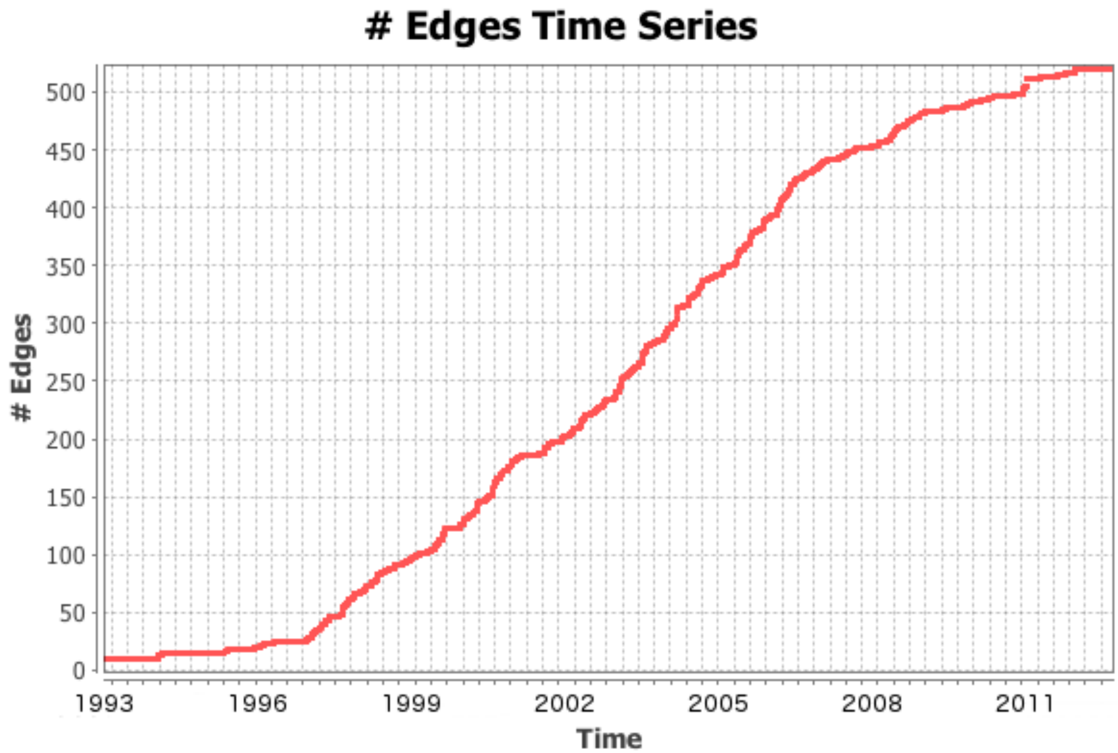

Figure 5: Total number of edges over time

5. Interpretation

Contrary to our original hypothesis that GLOBALink membership patterns would be different than other online communities given its double referral system, its membership count conforms to similar patterns of growth found in studies of other online communities (e.g. Backstrom et al. 2006; Firth et al. 2006) that adhere to the diffusion model, represented by the logistic curve (Figure 4). The total edge count follows the node graph in a parallel fashion. Both curves continue growing slowly until the late 1990's, followed by rapid growth through the early and mid 2000's. In the first three years of activity, there is no clustering at all (see Figure 6), followed by sharp spikes and volatile activity in the late 1990's, before reaching equilibrium around 2005-2006. The pattern also follows the node and edge growth, and the membership spike that occurs. This is expected, as there are too few nodes at the start to form separate and distinct clusters. The groups that make up the final modularity clusters begin taking shape around the year 2000, which is when we see a distinction between core and peripheral groups.

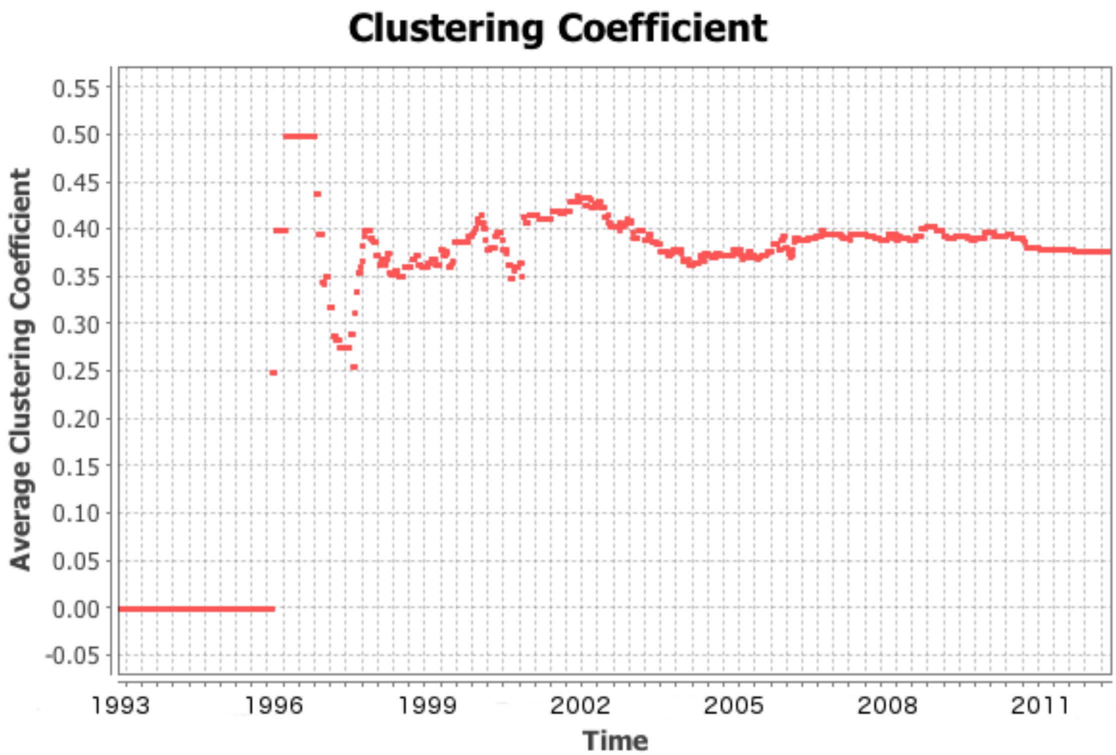

Figure 6: Clustering coefficient over time

Interestingly, the dates of the key shifts identified in the data match major events in the tobacco control community. For example, throughout much of the 1990's, international tobacco control was defined by individual country-to-country policy transfer, without much transnational discussion. Alternatively, the spike in the network growth occurring in the late 1990's parallels the launch of the FCTC negotiations and the large international investments made at that time. Regional negotiating blocks – increasingly visible in the network visualization beginning in the very early 2000's – were also a key feature of the FCTC negotiations. Toward the late 2000's, there is a slow decline in growth, indicating some saturation of the community and reduced international attention to the issue.

5.1. Modularity

Modularity is a measurement of how well a network can be divided into smaller clusters, or modules (Newman and Girvan 2004), and is useful in finding community structure (Newman 2006). High modularity indicates that a network has a higher rate of intra-module edges relative to inter-module ones. Gephi applies a modularity algorithm called the Louvain method, developed by Blondel and colleagues (2008) to find communities in the network. The resulting 0.305 value is not particularly high, but the 10 communities so identified had clear distinguishing characteristics.

Table 2: Modularity clusters

Module color	Highest referrer	Other prominent Members	Characteristic
Dark green	Australia	Philippines, Thailand, Vietnam	SE Asia
Light pink	Chile	Ecuador, Argentina, Brazil	South America
Purple	Samoa	Micronesia	Polynesia
Light blue	Oman	Saudi Arabia, Yemen, Egypt	Middle East
Dark blue	Canada	France, Congo, Algeria	French influence
Light yellow	Belgium	Portugal, Romania	Western Europe
Red	US	Switzerland, Hungary, Pakistan	
Light green	Guyana	Barbados, Trinidad & Tobago	Caribbean
Dark yellow	Denmark	Uzbekistan, Russian Federation, Poland	Eastern Europe
Dark pink	Honduras	Costa Rica	Central America

Each module is briefly described in Table 2. The referral network appears to create a majority of sub-groups that are divided geographically. These regions align closely with regional WHO offices that provided the platform for regional consultations and the formation of regional negotiating positions in between the global negotiating sessions held between 2000 and 2003. Other clusters appear to encompass the sphere of influence or language affinity between particular countries (e.g. Dark blue countries associate with French language/influence).

In the final static network graph (Figure 2), the United States is depicted as the largest node based on out-degree (89, Switzerland is 2nd with 66), which represents the highest amount of referrals to different countries, and leads by several other centrality measures. However, when the out-degree edges are weighted by the raw referral count, Switzerland is higher than the US by a significant margin, 499 to 274. The US also has the highest overall degree centrality (122, Switzerland is ranked 2nd with 76) and the highest betweenness (US, 0.21; Switzerland, 0.07). By most network metrics, the United States would seem to represent a very central node of high importance in this network. However, we expand our view to combine dynamic data with our static measurements to tell a different story.We see that in the early years, there was little growth in GLOBALink membership and activity was localized between several countries and Switzerland, headquarters of GLOBALink. In this sense, Switzerland represents the GLOBALink staff itself reaching out and actively helping others join in order to build the network. Indeed, throughout the 1990's and early 2000's, the GLOBALink director was very active in seeking out new members, attending dozens of tobacco control-related conferences around the world, and signing people up on the site. It is not until the late 1990's that we can distinguish local clusters being led primarily by Australia and Denmark. This corresponds with the start of the global tobacco control movement at WHO, where Australia acted as a strong leader throughout the FCTC negotiations, especially in the Western Pacific Region.

The case of the United States also requires the full longitudinal view of the network to understand its impact. The US has had members in GLOBALink from the very beginning of the community – as it was initially developed by the American Cancer Society. However, once GLOBALink is moved to the UICC (located in Geneva) and as the network really begins to grow in the 1990's, US referral activity was rather limited. Initial growth of the network during that period is led largely by Switzerland, and to a lesser extent, Australia, Canada, France, and the United Kingdom. But it is not until the mid 2000's that the US finally begins to assert influence on the network with a rapid increase in referrals. In terms of global tobacco control, this largely reflects the rather low-key role that US-based tobacco control activists played during the early days of the transnational tobacco control networking and the FCTC negotiations. The rapid increase in US-based networking interestingly coincides with the launch of the Bloomberg Initiative in 2006 and the influx of $500 million USD largely given to US-based institutions to administer in support of tobacco control efforts in select low- and middle-income countries. From our analysis, it appears that the Bloomberg investment got the US tobacco control community to engage more broadly in the global effort. However, by this point, most of the members being referred by the US belong to countries that already have some presence in the GLOBALink community. Thus instead of further extension and growth of the network, we witness the consolidation and growth of the US' centrality in the network.

Overall, the video exposes several important details in the network that are not immediately visible in the final static snapshot. When the network was first formed, only five of the ten modular-specific referral leaders were involved in the community. The final five join between the late 1990's (Chile) to the mid 2000's (Guyana). Not surprisingly, the first five countries are in the final core of the network with high degree centralities, while the later five are scattered in the periphery with low global degree centralities. We are also better able to understand the clustering coefficient graph.

5.2. Limitations

One limitation of this study is that individuals are not always defined by their reported country. For example, GLOBALink staff members were clearly identified in our study as Switzerland, obscuring the role of the network creators played in the network's early development, as opposed to the role that Switzerland itself plays in the global tobacco control community. Additionally, we limit our unit of analysis only to countries, as it was the only major attribute that members of GLOBALink can use to identify their affiliation. While other possible forms of affiliations (e.g. NGO's) might play important roles in the community, additional information was not required to be included by the members.

6. Future directions

This study provides an example of how dynamic network data can be analyzed in relation to real-world events. For example, changes in modularity coincided with an increase in GLOBALink membership during the period of the FCTC negotiation and the US emerged as a global leader in GLOBALink referrals after Bloomberg monies became available to combat global tobacco use. There are numerous areas for future research in relation to the policy implications, including further analysis of the impact of funding on the network evolution. While we could clearly see the influence of the US, more analysis is needed in regards to the other recipient countries. Follow-ups can also be done in regards to negotiating blocks and alliances that formed through the negotiations and whether that translated into lasting inter-country relationships. Future research will include more analyses regarding network factors associated with treaty ratification and implementation of related policies.

Research on the diffusion of policy innovations has often inferred contagion mechanisms in their adoption (Walker 1969; Gray 1973). Dynamic network visualizations and their associated metrics lend more insight into how networks and behaviors co-evolve. This type of analysis can help inform strategies needed to accelerate policy diffusion and/or improve discussions between opponents across international boundaries.

7. Conclusion

Data visualizations have been useful in many scientific fields, including social network analysis. The evolution of the sociogram, especially in conjunction with modern computers and software, has helped advance SNA studies, providing researchers with a better understanding of network characteristics such as structural patterns, positions that actors occupy, or where clusters emerge. As we push progress in studying longitudinal networks, the tools and methods used must also continue to evolve. Dynamic network analysis began as a collection of SNA-derived extensions, but its application in longitudinal network studies (e.g. terrorist cells, diffusion models) has demonstrated its necessity and utility. Research in how to conduct dynamic network studies, and consequently, the visualizations used, will help provide us with the necessary tools to better understand the unique nature of how networks are formed, patterns of evolution, and the metrics used to study them.

In this paper, we provided an example of how to use network visualizations to explore the dynamics of an online tobacco control community, GLOBALink. In particular, we focused on how a dynamic visualization can supplement limited static data. Through our dynamic network visualization of GLOBALink, for example, we have been able to visualize how and when the network grew, and how its evolution reflected major political and economic changes in the global tobacco control environment. Additionally, we showed that a static post-hoc sociogram might limit the contextual understanding in the history of how a network was formed, developed, and evolved. For example, while the sociogram representing GLOBALink initially showed the importance of Switzerland and the United States, visually and supported by classic network measurements, the longitudinal visualization (i.e. movie) exposed the full history of the network, revealing Switzerland's highly beneficial activities that helped shape the early network, with limited US involvement until late.

The dynamic graphical analysis, combined with geographic and network metric information, yielded insights into this diffusion process not discernible from a static analysis. The analysis also connected network analysis with real-world events. Advances in visualization methods should continue to be developed that can help researchers better explore network data, provide broader temporal context, and accurately connect specific events with network measurements. Ultimately, this research can inform international efforts to improve online community building and understand the influence of political and economic decisions have on the ways in which people interact and communicate with each other.

8. References

Backstrom, L., D. Huttenlocher, J. Kleinberg and X. Lan (2006). Group Formation in Large Social Networks: Membership, Growth, and Evolution. Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM.

Barabasi, A. L., H. Jeong, Z. Neda, E. Ravasz, A. Schubert and T. Vicsek (2002). "Evolution of the social network of scientific collaborations." Physica A: Statistical Mechanics and its Applications **311**(3): 590-614.

Barnett, G. A. (2001). "A longitudinal analysis of the international telecommunication network, 1978-1996." American Behavior Scientist **44**(10): 1638-1655.

Bastian, M., S. Heymann and M. Jacomy (2009). Gephi: An open source software for exploring and manipulating networks. International AAAI Conference on Weblogs and Social Media.

Bender-deMoll, S. and D. A. McFarland (2006). "The Art and Science of Dynamic Network Visualization." Journal of Social Structure **7**(2): 1-38.

Berger-Wolf, T. and J. Saia (2006). A framework for analysis of dynamic social networks. KDD '06: Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM Press.

Blondel, V. D., J.-L. Guillaume, R. Lambiotte and E. Lefebvre (2008). "Fast unfolding of communities in large networks." Journal of Statistical Mechanics: Theory and Experiment **10**.

Borgatti, S. and M. Everett (2000). "Models of Core/Periphery Structures." Social Networks **21**(4): 375-395.

boyd, d. (2006). "Friends, Friendsters, and MySpace Top 8: Writing Community Into Being on Social Network Sites." First Monday **11**(12).

boyd, d., S. Golder and G. Lotan (2010). Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter. Proceedings of the 43rd Hawaii International Conference on System Sciences, Hawaii.

Brandes, U. and C. Pich (2012). "Explorative Visualization of Citation Patterns in Social Network Research." Journal of Social Structure **12**(8).

Carley, K. (2003). Dynamic Network Analysis. Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers. K.

Carley, R. Breiger and P. Pattison. Washington, DC, The Nat. Acad. Press**:** 133-145.
Carley, K., M. Dombroski, M. Tsvetovat, J. Reminga and N. Kamneve (2003). Destabilizing Dynamic Covert Networks. Proceedings of the 8 International Command and Control Research and Technology Symposium.
Christakis, N. A. and J. H. Fowler (2010). "Social Network Measures for Early Detection of Contagious Outbreaks." PloS one **5**(9): e12948.
Chu, K.-H. and D. Suthers (2013). Technology-Mediated Relationships in a Socio-Technical System. Proceedings of the 46th Hawaii International Conference on System Sciences (HICSS 2013), Maui, Hawaii.
Correa, C. D. and K.-L. Ma (2011). Visualizing Social Networks. Social Network Data Analytics, Springer**:** 307-326.
de Nooy, W., A. Mrvar and V. Batagelj (2005). Exploratory Social Network Analysis with Pajek (Structural Analysis in the Social Sciences), {Cambridge University Press}.
Ellison, N., C. Steinfield and C. Lampe (2007). "The Benefits of Facebook "Friends:" Social Capital and College Students Use of Online Social Network Sites." Journal of Computer-Mediated Communication **12**(4): 1143-1168.
Falkowski, T., J. Bartelheimer and M. Spiliopoulou (2006). Mining and Visualizing the Evolution of Subgroups in Social Networks. IEEE/WIC/ACM International Conference on Web Intelligence, IEEE Computer Society.
Firth, D. R., C. Lawrence and S. F. Clouse (2006). "Predicting Internet-based Online Community Size and Time to Peak Membership Using the Bass Model of New Product Growth." Interdisciplinary Journal of Information, Knowledge, and Management **1**(1): 1-12.
Freeman, L. C. (2000). "Visualizing Social Networks." Journal of Social Structure **1**(1).
Gloor, P. A. and Y. Zhao (2004). TeCFlow - A temporal communication flow visualizer for social network analysis. ACM CSCW Workshop on Social Networks, Chicago.
Gray, V. (1973). "Innovation in the states: A diffusion study." The American Political Science Review **67**(4): 1174-1185.
Hether, H. J., S. Murphy and T. W. Valente (in-press). "It is better to give than receive: The role of social support, trust, and participation on health-related social networking sites." Journal of Health Communication.
Kairam, S., D. J. Wang and J. Leskovec (2012). The Life and Death of Online Groups: Predicting Group Growth and Longevity. In

Proceedings of the fifth ACM International Conference on Web Search and Data Mining, ACM.
Kossinets, G. and D. Watts (2006). "Empirical Analysis of an Evolving Social Network." Science **311**(5757): 88-90.
Krebs, V. E. (2002). "Mapping Networks of Terrorist Cells." Connections **24**(3): 43-52.
Levina, N. and E. Vaast (2005). "The Emergence of Boundary Spanning Competence in Practice: Implications for Implementation and Use of Information Systems." MIS Quarterly **29**(2): 335-363.
Liben-Nowell, D. and J. Kleinberg (2006). "The Link-Prediction Problem for Social Networks." Journal of the American Society for Information Science and Technology **58**(7): 1019-1031.
McCulloh, I. and K. Carley (2011). "Detecting Change in Longitudinal Social Networks." Journal of Social Structure **12**(3).
Mislove, A., M. Massimiliano, K. P. Gummadi, P. Drushel and B. Bhattacharjee (2007). Measurement and Analysis of Online Social Networks. In Proceedings of the 7th ACM SIGCOMM conference on Internet Measurement, ACM.
Moody, J., D. McFarland and S. Bender-deMoll (2005). "Dynamic Network Visualization." American Journal of Sociology **110**(4): 1206-1241.
Moody, J. and D. R. White (2003). "Structural cohesion and embeddedness: A hierarchical concept of social groups." American Sociological Review: 103-127.
Moreno, J. L. (1932). Application of the Group Method to Classification. New York, National committee on prisons and prison labor.
Newman, M. E. (2006). "Modularity and community structure in networks." Proceedings of the National Academy of Sciences **103**(23): 8577-8582.
Newman, M. E. and M. Girvan (2004). "Finding and evaluating community structure in networks." Physical review E **69**(2).
O'Madadhain, J., D. Fisher, S. White and Y.-B. Boey (2003). The JUNG (Java Universal Network/Graph) Framework, University of California, Irvine, California.
Perer, A. and B. Shneiderman (2008). Integrating statistics and visualization: case studies of gaining clarity during exploratory data analysis. Twenty-sixth annual SIGCHI conference on human factors in computing systems, ACM.
Rogers, E. (2003). Diffusion of Innovations, 5th Edition. New York, Free Press.
Stefano, A., D. Alonso and M. Pascual (2008). "A general model for food web structure." Science **320**(5876): 658-661.

Suh, B., L. Hong, P. Pirolli and E. Chi (2010). Want to be Retweeted? Large Scale Analytics on Factors Impacting Retweet in Twitter Network. 2010 IEEE Second International Conference on Social Computing (SocialCom), IEEE.
Susarla, A., J.-H. Oh and Y. Tan (2012). "Social Networks and the Diffusion of User-Generated Content: Evidence from YouTube." Information Systems Research **23**(1): 23-41.
Valente, T. W. (1995). Network models of the diffusion of innovations. Cresskill, NJ, Hampton Press.
Valente, T. W. (2005). Models and methods for studying the innovation of diffusion. Models and Methods in Social Network Analysis. P. J. Carrington, J. Scott and S. Wasserman. Cambridge, UK, Cambridge University Press**:** 98-116.
Valente, T. W. (2010). Social Networks and Health. Oxford, Oxford University Press.
Walker, J. L. (1969). "The diffusion of innovations among the American states." The American Political Science Review **63**(3): 880-899.
Waller, M. and S. Lipponen (1997). Smokefree Europe: A Forum for Networks, Finnish Centre for Health Promotion.
Wipfli, H., K. Fujimoto and T. W. Valente (2010). "Global Tobacco Control Diffusion: The Case of the Framework Convention on Tobacco Control." American Journal of Public Health **100**(7): 1260.