Joint Bachelor of Data Science
Bachelor End Project

Visualizing Dynamic Network structure by interactively slicing data on time-dimension

Yichen Wang
y.wang12@student.tue.nl

Supervisors:
prof. Roger Leenders
prof. Claudia Zucca

Eindhoven, June 2021

Acknowledgement

Abstract

Text

Table of Contents

1. Introduction

In recent years, the study of networks and the corresponding network analysis has become a topic of great interest (Moody, McFarland & Bender-DeMoll, 2005). By depicting the relationships among entities and analyzing the structures that arise from the recurrence of relations, the network analysis can be applied to a variety of fields such as social sciences, communication behavior, life sciences to help us in deep understanding of relationships and give better explanations of social phenomena (Chiesi, 2001).

Since Sociograms were introduced by Moreno (Moreno, 1953), visualization has long been an essential tool for analyzing social networks. Typically, the relevant research data was collected at a particular point in time, and researchers were using nodes and edges to represent objects and relationships with a static structure in which nodes never crash and edges maintain operational status forever. However, social networks evolve asynchronously, it has been shown that the structure of the network, as well as the characteristic features of entities, may not only mutually depend at a certain point in time, but also influenced from the past in a continuous-time dimension with the Markov process (Snijders, 2005). Thus, if we would like to observe the change and properties of the network over time, a continuous-time dimension would be introduced into the network, in which data is continuously observed and collected. Such continuous-time networks have dynamic structures that vary over time which nodes may come and go, edges may crash and recover (Rajaraman, 2006).

However, continuous-time network visualization has been a challenging subject for years, developers only have progressed for optimal network layout (Freeman, 2000; Brandes, Raab & Wagner, 2001) but are struggling with how to present the time-based connected data with an intuitive way to expose the information which folded in the timeline (Shi, Wang & Wen, 2011). It stems from the fact that the presentation of a relational structure of a network that consists of nodes and edges takes at least two dimensions, this leaves no effective space for visualization to represent the information from the timeline (Moody et al., 2005). Some tools for dynamic network visualization already exist though (Liu et al., 2015; Shi et al., 2015), they are not very sophisticated in representing the change of network structure over time, due to these tools disaggregate continuous-time so that the user's observation of the network can only be focused on a specific selected discrete point in time. Some other tools intend to overcome the limitations of high dimensionality. They slice the data equally in the time dimension, represent each slice as a separate network, and then present all the networks simultaneously on the interface with additional edges concatenating the same entities on different time slices (Everton, 2004). Nevertheless, these approaches require a certain level of background knowledge and reading ability from users, otherwise, users cannot clearly compare networks in different time slices and get informative results. As a temporary alternative, using the animation approach to display network movies by temporal views was applied in some continuous-time network studies, and some well-known visualization libraries, like *ndtv*, *vignette igraph*, they already could either provide animation features or call the interface of *animation* library to achieve such function. However, the animation approach has been questioned that it may not be very effective for network analysis tasks. The results of some experiments showed that static images provide better performance on readability than their animated counterparts (Farrugia & Quigley, 2011). In addition, each industry has different needs and segments for the research of the network and the functions of the corresponding visualization tools in this field. Whereas the approaches mentioned above do not carefully examine the meaning and implications of time in the formation of social networks, and the tools are not tailored to certain domains or to meet the needs at the application level.

The purpose of visualization is to have visual summary of information and easily identify patterns and trends. Thus, the visualization of continuous-time networks should not be limited to the traditional graph of nodes and edges, in other words, as long as it can help augment theoretical intuition (Moody et al., 2005), the visualization application of continuous-time networks can even be composed of multiple visualization components with different types of plots to present the information required by the user from multiple aspects and dimensions, starting from the research needs and user demands of a specific field.

Therefore, we propose a method to solve the problem of visualizing continuous-time networks in this paper. The method starts from the analysis of the characteristics of a representative dataset and the needs of its related research area, to examine carefully questions about the meaning of time in the formation of social networks and ensure the data are processed appropriately. Next, the data structure with class diagram are demonstrated to show the underline logic of data storage and manipulation. Afterwards, some important algorithms and implementations of graph theory and social network analysis are introduced that justify the type of questions that can be answered by our approaches which meet the needs of continuous-time network research for both the summary statistics over time and the changes of the network structure in the selected time range. Finally, the architecture of application and the design of GUI are presented to show how interactive operation proceeded between multiple visualization components.

2. Dataset and Prior Research

2.1 Dataset

!! *UNDER CONSTRUCTION* !!

The dataset was collected from a primary school in Lyon, France over two days of school activity (1^{st} Oct and 2^{nd} Oct 2009). The researcher observed 232 children and 10 teachers wearing wearable proximity sensor to record the face-to-face proximity contacts with a 20-second temporal resolution among individuals (Gemmetto, Barrat & Cattuto, 2014). The dataset contains a total of 242 agents with characteristics such as id, grade class, gender, and 125772 records of contacts with time, both agents of contacts, and their class.

The researchers aimed to analyze this continuous-time data to assess the mode and severity of transmission of infectious diseases in schools and to make recommendations on how schools should respond to the risk of major infectious diseases. Therefore, this data is both challenging and relevant to the current social situation concerning the prevention and control of the Covid-19 pandemic.

2.2 Objectives and Demands

3. Methodology

3.1 Data Structure

3.2 Theory and Algorithms

3.3 Application Architecture

3.4 GUI design

4. Results

5. Discussion

6. Conclusion

Text

7. Recommendations

References

Brandes, U., Raab, J. & Wagner, D. (2001). Exploratory Network Visualization : Simultaneous Display of Actor Status and Connections. *Journal of Social Structure*. 1

Chiesi, A. (2001). Network analysis. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social behavioral sciences* (p. 10499-10502). Oxford: Pergamon. Retrieved from https://www.sciencedirect.com/science/article/pii/B008043076704211X doi: https://doi.org/10.1016/B0-08-043076-7/04211-X 1

Everton, S. (2004, 01). A guide for the visually perplexed: Visually representing social networks.
1

Farrugia, M. & Quigley, A. (2011). Effective temporal graph layout: A comparative study of animation versus static display methods. *Information Visualization*. doi: 10.1057/ivs.2010.10 1

Freeman, L. C. (2000). Visualizing Social Groups. In *Journal of social structure.* 1

Gemmetto, V., Barrat, A. & Cattuto, C. (2014). Mitigation of infectious disease at school: Targeted class closure vs school closure. *BMC Infectious Diseases*. doi: 10.1186/s12879-014-0695-9 3

Liu, Q., Hu, Y., Shi, L., Mu, X., Zhang, Y. & Tang, J. (2015). EgoNetCloud: Event-based egocentric dynamic network visualization. In *2015 ieee conference on visual analytics science and technology, vast 2015 - proceedings.* doi: 10.1109/VAST.2015.7347632 1

Moody, J., McFarland, D. & Bender-DeMoll, S. (2005). *Dynamic network visualization.* doi: 10.1086/421509 1, 2

Moreno, J. L. (1953). *Who shall survive? Foundations of sociometry, group psychotherapy and socio-drama.* 1

Rajaraman, R. (2006). *Introduction to Dynamic Networks Models, Algorithms, and Analysis* (Tech. Rep.). Retrieved from www.ccs.neu.edu/home/rraj/Talks/DynamicNetworks/DYNAMO/ 1

Shi, L., Wang, C. & Wen, Z. (2011). Dynamic network visualization in 1.5D. In *Ieee pacific visualization symposium 2011, pacificvis 2011 - proceedings.* doi: 10.1109/PACIFICVIS.2011.5742388 1

Shi, L., Wang, C., Wen, Z., Qu, H., Lin, C. & Liao, Q. (2015). 1.5d egocentric dynamic network visualization. *IEEE Transactions on Visualization and Computer Graphics*, *21*(5), 624-637. doi: 10.1109/TVCG.2014.2383380 1

Snijders, T. (2005). Models for longitudinal network data. *Models and Methods in Social Network Analysis*. 1

A. My First Appendix

In this file (Appendices/Appendix_A.tex) you can add appendix chapters, just as you did in the Document.tex file for the 'normal' chapters. You can also choose to include everything in this single file, whatever you prefer.