Final Project Proposal

Yi Chen

Teachers College Columbia University

Multidimensional Scaling, Clustering, and Network Models

Social network analysis refers to a collection of technique methods for describe the interactions among the individuals and to represent their relationship. In this final project, I will explore the latent space network model (Sweet and Adhikari, 2020) for social inference. The uniqueness of latent space network model that it is able to identifying the latent space location which could be used for explain the observed network structure.

In this model frame work, the social network structure by a adjacency / tie matrix,

, where is the value of the edge from actor to actor , which is usually binary value to represent the ties absent or present. The general latent space model that Sweet and Adhikari (2020) proposed for binary cross-sectional network data can be expressed as follow:

, where is the adjacent matrix, is the position of individual in the -dimensional latent space, which is sampled from a multidimensional normal distribution with and set as the mean and variance hyperparameter, and is the intercept term in the logistic regression which is also sampled from the normal distribution. The observed graphic structure dependent on the latent space . is the outcome variable measured at time and for individual . is the G nearest neighbor (defined based on the latent space location) for individual , and is their relative weight of neighboring . Consequently, the outcome of individual at time point can be explained by the intercept term , the outcome at , and weighted average of the outcome at for the G closeted neighbors in the latent space. Based on this model framework, a social network (taking the undirected network as example) depend on the dimensional latent space location. Consequently, the proximity of individual in the latent space can be used to represent their relationship in the network space. This framework can be estimated in a fully Bayesian approach (MCMC) using Rjags or Stan.

This final project aims at three main takes: (1) realize the Bayesian estimation program for latent space model, (2) provide the simulation study with different group size, group interaction density level, and whether or not the group has a clear leader.

**The idea for the future research**

However, the framework of this latent space network model has some limitations. First, the latent space parameter is hard to interpret psychologically seven the latent space is able capture the general relationship in the social network. Second, the latent space variables and the observed network structure are fixed over the time. The only time-varying in this model is the individual outcome. Finally, the outcome variables are defined in the individual level. Thus, the interactions between the individual could not be used to explain the group collaboration and team performance. To fixed these issues, I redesign the model for the purpose of measuring the dynamic group collaboration with the evidence from the social network interactions.

In this model framework, each time point , we will have adjacent matrix . The time point we defined here is not the instant, but a range of time. For example, if our research is about measuring the group collaboration of the same group of students over a whole semester. The time point is each class during the semester. Consequently, the interactions among the students during the class could be collected. Similarly, the interaction between any pair of students in the team depend on the latent space variable of each individual and the fixed intercept (the overall class interactive level).

Different from Sweet and Adhikari (2020), the individual latent variables at time point depend on their previous position in the latent space.

However, I have to admit that fixing the variance term may not reasonable since the variability of the latent space variable may not consistent. Finally, the team outcome (e.g., the group overall performance) at time point can be described as:

, where we use the data reduction function to reduce the dimension of from into . The possible function could the mean of the individual position on dimension Then, is a dimensional vector, with each element represent the relative importance of dimension in predict the team outcome in the next time point.

References

Sweet, T. and Adhikari S. (2020). A Latent Space Network Model for Social Influence. *Psychometrika*, https://doi.org/10.1007/s11336-020-09700-x.