Effects of Network Structure on Subjective Preference Diversity

ANNE LIN, University of Michigan Ann Arbor
ANDRÉS ABELIUK AND EMILIO FERRARA, Information Science Institute, University of Southern California

1. INTRODUCTION

As we spend more time online, we are increasingly exposed to and influenced by the taste preferences and attitudes of others. However, the information can come in different forms. For example, in the case of search engine results or billboard music rankings, people take in the aggregate signal of others. A structurally similar situation is when people take in information from highly connected communities, such as subreddits. At other times, in the case of Instagram feed or friends Spotify playlists, people take in the signals of individuals that they are connected to. As institutions and organizations increasingly place more emphasis on cultivating diversity perspectives, it is important to understand the structural conditions under which diversity occurs. In this study, we examine how network connectivity affects the diversity of subjective preference.

2. RELATED WORKS

Our study is most closely related to past studies on the topic of ”social learning” [Miller and Dollard 1941], which looks at how network structure affects the process of group problem solving. The majority of such research have an experimental design where agents are tasked with solving a problem that has an objectively correct answer [Golub and Jackson 2010, Becker et al. 2017, Shirado and Christakis 2017]. Our experiment differs from traditional social learning literature as it focuses on a subjective task where there is no pressure to conform based on the assumption that others have expert information that one does not.

The research that looks at the spread of subjective information in a network can be split into two groups: the creation of social norms [Centola and Baronchelli 2015, Latané and L'Herrou 1996] and information cascades [Bond et al. 2012, Bikhchandani et al. 1992]. Research on the creation of social norms differs from our experiment in that the agents are explicitly rewarded for converging their solution. This could be for monetary rewards [Centola and Baronchelli 2015] or for the sake of efficiency or to make oneself more likable [Chartrand and Bargh 1999, Cialdini and Goldstein 2004]. Our study differs from the studies on social norms in that it shows that homogeneity in subjective preferences can arise even if there is no social pressure to have the same tastes or perform the same actions.

Research on information cascades, typically focuses on how one source of information spreads rather than the difference of information sources amongst a group [Guille et al. 2013]. By creating an agent-based model to study the relationship between the adoption of competing information and network structure, we are able to examine how change in network structure affects the process of information adoption as well as the dynamic of the different information and how they interact with one another.

3. METHODS

Our model comprises of an undirected social network where each node represents an agent, and each edge represents a mutual connection between the agents. At each time step, each agent in the network is presented with an item, which they either ”like” or ”dislike”. At the first time step, each agent is presented with a random item to evaluate. Subsequently, the item presented to an agent at each time step is the item that is the most-liked among the other agents that it is linked to. Each item has a quality score that corresponds to the probability of an agent "liking" the item. The distribution of quality scores is uniform across all items.

In each condition, there are 100 agents in a network. We test out a set of different numbers of items (50, 100, 150, 200 and 250) across all conditions. For each value, the agents evaluate 1/10 the number of total items. We run the model 25 times for each different variable value. We test out three conditions:

1) *Independent condition*: There are no edges in the graph, thus each agent rates their items independently, with no influence from other agents.
2) *Social influence condition:* Each agent is placed within a network, where each agent is linked to 10 other agents. We use a watts-strogatz small-world network [Watts and Strogatz 1998] and vary the rewiring probability.
3) *Aggregate condition*: The agents are placed within a complete network, where each agent is linked to all the other agents.

We use two measurements to analyze the outcome:

1) Group opinion diversity

To measure this, we calculate the heterogeneity score of the group of agents across the conditions. We calculate the heterogeneity score by averaging the difference between a pair of liked-items list and a pair of disliked-items list, and then normalizing the number by dividing it by the number of trials. We calculate the score for all pair-wise combination of agents and finally calculate the average of the scores. At a high level, this score represents the percentage of non-over lapping items in two agent's liked and disliked items lists. Therefore a heterogeneity score of 1 represents two agents not having any overlap in their outcomes; a score of 0 represents two agents having the exact same outcome.

$$\sum avg\left(\frac{l_1 \cap l_2}{t}, \frac{d_1 \cap d_2}{t}\right)$$

In the heterogeneity equation above, l represents a list of liked items, d represents a list of disliked items, and t represents the number of trials in the model run. It is similar to the Jaccard index, with a different denominator, which takes into account the weight of the liked/disliked lists in respect to their lengths.

2) Item popularity distribution

To measure this, we calculate the popularity score of each item after a model run. For the popularity score, each item's score increases by one with each agent liking it, and decreases by one with each agent disliking it. We then calculate the distribution of the items using the gini index, which was first developed to measure a population's income inequality. We adapted Raffinetti and Aimar's GiniWegNeg algorithm [Raffinetti and Aimar 2016] to account for the negative popularity scores, since the traditional gini algorithm does not take negative values as input.

4. RESULTS

1) Group opinion diversity

We hypothesize that the less connected people are to each other within a social network, the more opinion diversity the group will have. Our findings match our hypothesis: the independent condition, where the group had no connections at all, has the highest heterogeneity score; the aggregate condition, where every agent is connected with each other, has the lowest heterogeneity score. In order to test for significance, we conducted a one-way ANOVA test. The aggregate condition's heterogeneity score was significantly lower than the social influence condition heterogeneity scores when there are 50 items ($F(2, 297) = 31.95$ $p < 0.001$) and 100 items in the model ($F(2,297) = 5.81$, $p <$

0.05). Furthermore, as the number of items in the model increases, the diversity of the agents decreases (See Figure 1).

2) Item popularity distribution

We hypothesize that the more connected a network is, the more likely it is that the group's attention will be over-saturated by quality items that go viral. This is in line with what we find in the model runs. The independent condition produces the lowest gini index, meaning that the item popularity distribution is the least skewed; the aggregate condition produced the highest gini (See Figure 2). The gini indexes of the social influence condition are consistently significantly lower than the gini index of the aggregate condition across all numbers of items ($F(2, 297) = 77.75$, $F(2, 297) = 61.71$, $F(2, 297) = 382.40$, $F(2, 297) = 216.59$, $F(2, 297) = 77.75$ for items = 50, 100, 150, 200, 250, respectively. $p<0.01$ for all numbers of items).

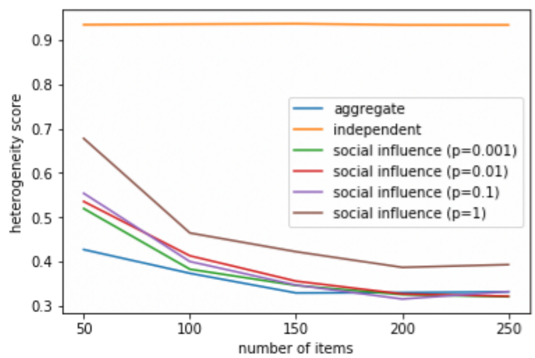

Fig. 1

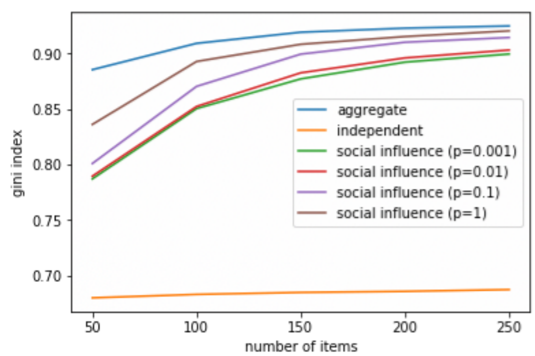

Fig. 2

5. DISCUSSION

Our study seems to point to a paradox between diversity and social connectedness. On one hand, in order for diverse ideas to spread quickly, an environment that facilitates close social connection is necessary. However, on the other hand, as this study shows, interaction between people in such an environment leads to more homogeneity in the group over time.

One limitation to this model is that agents choose to evaluate the most popular item amongst their social connections. In the future, we hope to adjust the model such that each agent probabilistically decides on which item to evaluate, where the more popular an item is, the more likely the agent is to evaluate it. Another interesting direction to explore is a model where agents have specialized roles. For example, there could be a model design where a small set of agents are "innovators" who are more likely to explore new options, while the rest are "imitators" who exploit known options, similar to the set up in Wisdom et al.'s social learning model [Wisdom et al. 2013].

In many cases, the adoption of information is a type of complex contagion that greatly depends on the network strength of the ties [Centola and Macy 2007]. In our model, we have simplified the network in assuming that all tie strengths are equal. However, in the real world, it is likely that an agent does not automatically evaluate the most popular resource amongst all their ties, but rather weight the resources based on the strength of their social connections. The goal of our model is to understand how network structure influences one's opinions independent of social influence. However, it would be necessary to incorporate previously studied factors of complex contagion and opinion convergence if the goal is to create a more accurate, predictive model.

REFERENCES

Becker, J., Brackbill, D. and Centola, D., 2017. Network dynamics of social influence in the wisdom of crowds. *Proceedings of the national academy of sciences*, *114*(26), pp.E5070-E5076.
Bikhchandani, S., Hirshleifer, D. and Welch, I., 1992. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, *100*(5), pp.992-1026.
Bond, R.M., Fariss, C.J., Jones, J.J., Kramer, A.D., Marlow, C., Settle, J.E. and Fowler, J.H., 2012. A 61-million-person experiment in social influence and political mobilization. *Nature*, *489*(7415), pp.295-298.
Centola, D. and Baronchelli, A., 2015. The spontaneous emergence of conventions: An experimental study of cultural evolution. *Proceedings of the National Academy of Sciences*, *112*(7), pp.1989-1994.
Centola, D. and Macy, M., 2007. Complex contagions and the weakness of long ties. *American journal of Sociology*, *113*(3), pp.702-734.
Chartrand, T.L. and Bargh, J.A., 1999. The chameleon effect: the perception–behavior link and social interaction. *Journal of personality and social psychology*, *76*(6), p.893.
Cialdini, R.B. and Goldstein, N.J., 2004. Social influence: Compliance and conformity. *Annu. Rev. Psychol.*, *55*, pp.591-621.
DeGroot, M.H., 1974. Reaching a consensus. *Journal of the American Statistical Association*, *69*(345), pp.118-121.
Golub, B. and Jackson, M.O., 2010. Naive learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics*, 2(1), pp.112-49.
Guille, A., Hacid, H., Favre, C. and Zighed, D.A., 2013. Information diffusion in online social networks: A survey. *ACM Sigmod Record*, *42*(2), pp.17-28.
Latané, B. and L'Herrou, T., 1996. Spatial clustering in the conformity game: Dynamic social impact in electronic groups. *Journal of personality and social psychology*, *70*(6), p.1218.
Miller, N.E. and Dollard, J., 1941. Social learning and imitation.
Raffinetti, E. and Aimar, F., 2016. GiniWegNeg.
Shirado, H. and Christakis, N.A., 2017. Locally noisy autonomous agents improve global human coordination in network experiments. *Nature*, *545*(7654), pp.370-374.
Watts, D.J. and Strogatz, S.H., 1998. Collective dynamics of 'small-world'networks. *nature*, *393*(6684), p.440.
Wisdom, T.N., Song, X. and Goldstone, R.L., 2013. Social learning strategies in networked groups. *Cognitive science*, *37*(8), pp.1383-1425.