Algorithmically Mediating Communication to Enhance Collective Decision-Making in Online Social Networks

JASON W. BURTON and ULRIKE HAHN, Birkbeck, University of London
ABDULLAH ALMAATOUQ, Massachusetts Institute of Technology
M. AMIN RAHIMIAN, University of Pittsburgh

1. INTRODUCTION

Many collective decision-making contexts involve communication among group members. Sometimes this communication helps the collective reach an accurate decision because it allows individuals to gain otherwise unknown information from their peers, but sometimes this communication gives rise to detrimental social influence or "groupthink." Whether communication is ultimately good or bad for a group's collective decision-making depends on the underlying network structure (i.e., who communicates with whom): high levels of connectivity and free-flowing information can lead to "excess correlation" (i.e., correlation between individuals that is not accuracy inducing) [Jönsson et al. 2015]; high levels of centralization can lead to certain individuals wielding excessive influence over the network [Becker et al. 2017]; and a lack of structural plasticity can prevent networks from effectively responding to feedback about individuals' performance [Almaatouq et al. 2020b]. Despite abundant knowledge on the relationship between network structure and collective accuracy, strategies for exploiting network structure to increase collective accuracy remain under-explored. In the present work, we experiment with one such strategy, *rewiring algorithms*, which mediate online communication by manipulating social networks' structure. Crucially, the algorithms considered may improve accuracy by modifying connectivity based on the distribution of participant responses alone, that is, without access to a ground truth on the issue at the time of communication.

2. METHOD

We built an online multiplayer experiment with the Empirica software [Almaatouq et al. 2020a] and recruited 704 participants aged 18–69 ($M = 34.28, SD = 9.87$) via Amazon's Mechanical Turk crowdsourcing platform. Participants were randomly assigned into 16-person networks in one of the four network treatments (static, mean-extreme, polarize, or scheduled) and tasked with a "Collaborative Prediction Game" that consisted of ten rounds with five stages each. Each round of the game involved predicting the probability of one near future event occurring in reality (e.g., "Bitcoin will be valued at less than $30,000 on 8 February 2021"), with participants first providing a probabilistic prediction and short rationale for their prediction independently, and then proceeding through four stages of communication where each participant would view the responses of their network neighbor(s) and revise their own prediction and rationale. A total of 44 networks completed the study (11 per treatment) and participants were given monetary incentives for collective accuracy.

2.1 Network Treatments

In the *static* network treatment, our control condition, participants communicated across static, unchanging networks. For our experimental network treatments, rewiring algorithms manipulated the networks' structure stage-by-stage on the basis of the individuals' responses. Specifically, we considered three such algorithms that we previously developed with agent-based simulations: a *mean-extreme* algorithm, a *polarize* algorithm, and a *scheduling* algorithm.

The mean-extreme algorithm aims to increase the average accuracy of individuals in a network by directing social influence towards individuals with potentially erroneous, outlying estimates. The algorithm first calculates the mean estimate in a network at a given time point and identifies which side of the scale midpoint (0.5 on a 0-1 probability scale) the network's mean estimate lies. If the network's mean estimate is less than the midpoint, the algorithm identifies the agent with the lowest estimate and adds directed, outgoing ties to the three agents with the highest estimates. If the network's mean estimate is greater than the midpoint, the algorithm identifies the agent with the highest estimate and adds directed, outgoing ties to the three agents with the lowest estimates. This procedure ultimately brings the estimates of outlying individuals closer to the mean.

The polarize algorithm aims to maintain the diversity of estimates in a network and prevent a potentially biasing homogenization. It first identifies the two most extreme agents on either side of the current distribution of estimates (i.e., the agent with the highest estimate and the agent with the lowest estimate) and cuts all incoming ties to these agents so as to preserve their beliefs from social influence. Then, the influence of these extreme agents is increased by granting each of them two directed, outgoing ties to "core" agents. These core agents are the four individuals with the median estimates in the network (e.g., in a 16-agent network, the agent with the lowest estimate receives an outgoing tie to the agents with the 7th and 8th lowest estimates, and the agent with the highest estimate receives two outgoing ties to the agents with the 9th and 10th lowest estimates). The net effect of this procedure is that the diversity of beliefs (measured as variance) is increased by ensuring both extreme, "polar" sides of the belief spectrum are heard.

The scheduling algorithm differs from the mean-extreme and polarize algorithms in that it prescribes (or "schedules") a network structure of intermixing dyads, irrespective of individuals' estimates. Specifically, the algorithm pairs agents at each time point such that no agent speaks to the same agent twice, but each individual will have the opportunity to be in possession of all the available information in the network by the end of four rounds of communication. In this way, scheduled networks will have avoided any redundant interactions from taking place—each dyad at each time point will consist of two individuals sharing information received from individuals in the network that the other has not interacted with. This algorithmic approach offers an alternative for situations where access to individuals current estimates at each time point is not available.

3. RESULTS

Our analyses focus on the accuracy of the collective, mean responses of each network pre- and post-communication. In particular, we investigated the following three questions: (1) How did networks' average collective error squared (CES) differ between treatments? (2) How did communication affect CES within each network, between treatments? (3) How did the different rewiring algorithms influence networks' collective confidence calibration?

To address the first question, we used a linear mixed effect model with each groups' average collective error squared (CES) across all events predicted as the dependent variable, the network treatment as a fixed effect, and random intercepts by group. This analysis suggests that there is no significant effect of the rewiring algorithms on collective accuracy ($F(3, 436) = 0.784, p = 0.503$), meaning that, on average, networks to which a rewiring algorithm was applied did not produce more accurate predictions than static networks. However, this analysis does not account for certain key confounding variables—namely, the initial network structure and initial predictions in each network.

In addressing the second question we are able to side-step the potential confounding effects of initial network structure and initial predictions by evaluating the effect of communication *within* each

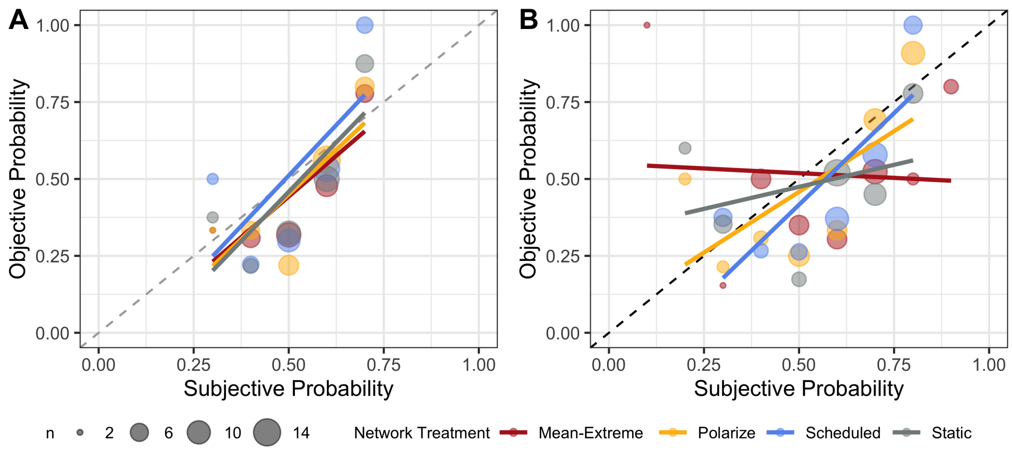

Fig. 1. Calibration of collective predictions. Predictions (i.e., subjective probabilities) are "binned" by rounding down to the nearest tenth decimal place, and the objective probability for each bin is calculated as the proportion of the events that occurred in reality. The dashed diagonal line represents perfectly calibrated predictions. More extreme subjective than objective probabilities represent overconfidence. Each point's size represents the number of events in the respective bin (n). Points where $n = 1$ have been excluded. **[A]** Pre-communication calibration. **[B]** Post-communication calibration.

network. That is, instead of directly comparing the accuracy of networks' collective predictions post-communication between treatments, we compare the *change in accuracy* between each network's prediction pre- and post-communication. Upon re-fitting our linear mixed effect model with networks' change in CES as the dependent variable, we find a significant treatment effect ($F(3, 436) = 2.722, p = 0.044$): networks mediated by the polarize algorithm experienced a decrease in CES following communication (change in CES, $M = -0.010, SE = 0.009$), whereas communication led to an increase of CES in mean-extreme (change in CES, $M = 0.025, SE = 0.009$) and static networks (change in CES, $M = 0.018, SE = 0.009$), and communication was neither beneficial nor detrimental to scheduled networks' CES (change in CES, $M = 0.000, SE = 0.009$).

Finally, we analyzed the calibration of networks' collective predictions pre- and post-communication. Calibration measures the extent to which subjective degrees of belief (probabilities) match objective probabilities [Fischhoff et al. 1977]. To assess calibration we "binned" the networks' predictions by rounding them down to the nearest tenth decimal place (e.g., 0.12 and 0.19 both become 0.1) and calculated the proportion of events in each bin that occurred. If a network's predictions are perfectly calibrated, the proportion of events in each bin that actually occurred match the the bin value. Fig. 1 shows collective calibration of networks in each treatment, pre-communication and post-communication, side-by-side. All networks were well-calibrated pre-communication, but communication affected calibration differently depending on the network treatment. In static networks it lead to overconfidence in collective predictions, while polarize and scheduling algorithms mitigated this effect, and the mean-extreme algorithm exacerbated it.

4. CONCLUSION

These results provide initial empirical evidence that mediating communication in social networks with different rewiring algorithms can influence the accuracy of collective decisions. Although we did not observe a statistically significant treatment effect on post-communication collective predictions, the findings that within-group effects of communication and collective calibration were influenced by the rewiring algorithms encourage further investigations.

REFERENCES

Abdullah Almaatouq, Joshua Becker, James P Houghton, Nicolas Paton, Duncan J Watts, and Mark E Whiting. 2020a. Empirica: a virtual lab for high-throughput macro-level experiments. *arXiv preprint arXiv:2006.11398* (2020).

Abdullah Almaatouq, Alejandro Noriega-Campero, Abdulrahman Alotaibi, PM Krafft, Mehdi Moussaid, and Alex Pentland. 2020b. Adaptive social networks promote the wisdom of crowds. *Proceedings of the National Academy of Sciences* 117, 21 (2020), 11379–11386.

Joshua Becker, Devon Brackbill, and Damon Centola. 2017. Network dynamics of social influence in the wisdom of crowds. *Proceedings of the National Academy of Sciences* 114, 26 (2017), E5070–E5076.

Baruch Fischhoff, Paul Slovic, and Sarah Lichtenstein. 1977. Knowing with certainty: The appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance* 3, 4 (1977), 552.

Martin L Jönsson, Ulrike Hahn, and Erik J Olsson. 2015. The kind of group you want to belong to: Effects of group structure on group accuracy. *Cognition* 142 (2015), 191–204.