NO-REGRET LEARNING ALGORITHM IN BAYESIAN PERSUASION GAMES

COMBATING MISINFORMATION USING GAME-THEORETIC INCENTIVES

A Thesis Prospectus
In STS 4500
Presented to
The Faculty of the
School of Engineering and Applied Science
University of Virginia
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Computer Science

By
Gustavo Moreira

November 1, 2021

Technical Team Members:
Jibang Wu

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

Signed: Gustavo Moreira   Date: November 1, 2021

ADVISORS

Catherine Baritaud, Department of Engineering and Society

Haifeng Xu, Department of Computer Science

A core concept in the study of multi-agent systems is that of *persuasion*, the notion that one agent's actions in a system can result in another agent choosing a different course of action then that which they would originally have chosen. This framework has a wide variety of applications in various social settings, including advertising, courts, lobbying, financial disclosure, and political campaigns, as each of these involves independent rational agents choosing whether to communicate information to each other in order to influence others' decisions (Kamenica & Gentzkow, 2009, p.3). This notion of persuasion also has applications in social coordination setting, specifically in organizing groups online and in-person into collective action (Ito, 2011, p.1). In economics, this descriptive framework has been called *Bayesian persuasion*, where the Bayesian term indicates that involved agents solve inference problems using Bayesian probability update rules to inform their decisions (Icard, 2018, p.79). Finding the optimal strategy for each agent in this generalized persuasion scenario would lend key insights on how to effectively communicate information to others in order to induce certain actions from them (Dughmi & Xu, 2019, p.1). However, this task proves to be exceedingly difficult in practice, and as such assumptions about the range of possible agent choices need to be made in order to make meaningful progress. More specifically, the general Bayesian persuasion scenario turns out to be very difficult to model mathematically due to the wide array of variables involved, including individual agents' private preferences, the state of the environment around the agent, and others (Liao, 2021, p.1).

The goal of the technical project is to design a generalized algorithm for a persuader (typically called a "Sender" in the literature) in a Bayesian persuasion setting to learn how to send signals to a "Receiver" such that the sender receives the highest reward possible (Kamenica & Gentzkow, 2009, p.1). While such algorithms exist in specific instances, the goal of this project will be to make the algorithm “no-regret”, meaning that it will have some guarantee of how optimal the reward is during the learning process rather than just converging on optimal policy eventually. The tightly coupled STS research project will use the insights discovered during the design of the aforementioned algorithm for the task of minimizing the spread of misinformation on social media platforms, leveraging existing research on collectivizing online action (Jasper, 2004, p.3). People are more likely to comply with intrinsically incentivized behaviors rather than externally enforced behaviors, especially if the external enforcement seems unmotivated. So, if social media users can be persuaded to verify sources for their posts, they are much more likely to do so than if they are simply removed from platforms for posting misinformation (Valenzuela, Halpern, Katz, & Miranda, 2019, p.3). The work on the technical project has already begun and will continue through April 2022, while the STS research paper will begin writing in February 2022.

NO-REGRET ALGORITHM IN BAYESIAN PERSUASION GAMES

McCloskey & Klamer (1995) said that “One quarter of the GDP is persuasion” (p.1), implying that a large portion of the economy depends on the understanding of how independent rational agents can induce actions in other independent rational agents. Such aspects of the economy include things like stock trading, investment banking, monetary policy, and more (Dughmi & Xu, 2019, p.2). However, persuasion has far broader applications than just economics, including political campaigns and collective action. Bayesian persuasion is merely a formal mathematical model used to analyze persuasion using the existing tools of algorithmic game theory (Kamenica & Gentzkow, 2009, p.3). Using Bayesian persuasion as a general model, one can look at specific instances of persuasion like the stock market and use the specifics of that situation to add additional constraints to variables as needed. However, because the algorithm ought to be generally applicable, the goal of any learning algorithm for Bayesian persuasion scenarios should be to impose as few restrictions as possible.

In order to proceed with designing an efficient algorithm for calculating optimal strategies for a Bayesian persuasion scenario, the constraints on the mathematical formulation need to be formalized. It is assumed there is a single Sender and Receiver, and the primary goal of each agent is to maximize their utility according to some utility function, where the utility drawn uniformly from a set probability distribution. The Sender can observe some state of nature, an analogue for some information about the environment such as a buyer's belief about a stock, drawn from an unknown probability distribution, and using that information can send a signal to the Receiver, who cannot themselves observe the state of nature, in order to try to influence the Receiver's decision. The Receiver then takes an action, and the utility for both the Sender and the Receiver is entirely determined by the state of nature and the action of the Receiver; this entire process is then repeated for T rounds until the agents have converged on some policy (Kamenica & Gentzkow, 2009, p.7). This also implies that the Sender does not themselves control the utility they receive; rather, only their signal's ability to persuade the Receiver has an effect on the utility. However, in the modification of the traditional Bayesian persuasion scenario used for this project, it is assumed that the Receiver is always taking the action that is a "best response" to the strategy used by the Sender; it is assumed that given the signaling strategy of the Sender, the Receiver is choosing the action that will result in the highest utility for the Receiver. This modification drastically simplifies the problem because only the Sender needs to go through the learning process, while simultaneously not limiting the general applicability of any designed algorithm because the express goal was to calculate optimal Sender behavior rather than optimal Receiver behavior. This assumption about Receiver best response is common in the literature and does not limit the applicability of the results of this project severely (Arieli & Babichenko, 2019, p.1). The algorithm developed is said to be "no-regret" if the actions taken by the agent in the learning process have an average reward no worse than the reward the agent would have received if they had simply chosen some constant action a from the beginning (Greenwald, Jafari, Gondek & Ercal, 2001, p.1). More specifically, an algorithm is said to be no-regret if it satisfies the following equation:

$$R(T) = \frac{1}{T}\left[\sum_{t=1}^{T} c^t(a^t) - \sum_{t=1}^{T} c^t(a)\right] \tag{1}$$

where $c^t(a)$ is the cost being minimized for each action a and round t, and a^t is the action taken by the Sender at round t. In the Bayesian persuasion case, these actions correspond to probabilities over the possible signals the Sender could send to the Receiver. Because the reward for the Sender is essentially based solely on their signaling policy, the goal of the algorithm will be to find a signaling policy that optimizes the Sender's reward for each possible best response of the Receiver (Wei, Yu & Neely, 2020, p.1). The work for the development of this algorithm has already begun as part of the SIGMA Lab in the Computer Science Department at UVa, under the advisership of Haifeng Xu and Jibang Wu, a Ph.D. candidate. The algorithm once fully developed and implemented should be able to quickly determine a signaling strategy that is optimal for each possible Receiver action, and thus determine the maximum possible utility for the Sender. Once completed, the algorithm and related work will be submitted as a paper to a variety of computer science conferences such as ICML 2022. The description of the algorithm would then be applicable to a variety of scenarios by adding additional constraints to the Bayesian persuasion scenario based on the specific application.

COMBATING MISINFORMATION WITH GAME-THEORETIC INCENTIVES

The spread of misinformation on social media is a well-known phenomenon that is still very difficult to combat for a variety of reasons. Misinformation in this context is used to refer to specifically false or inaccurate information that is created with the explicit purpose of deception while being propagated both intentionally and unintentionally (Wu, Morstatter, Carley & Liu, 2019, p.1). Part of the reason that it is so difficult to prevent the spread of misinformation is because the incentive structures that exist for social media platforms prioritizes engagement rather than any sort of general social welfare, which leads to suboptimal outcomes for the users of the platforms despite the increased profit margins for the platform owners (Au, Ho & Chiu, 2021, p.5). A variety of different counter-measures for such incentive structures have been proposed, including forcing platforms to change policies via legislation and encouraging users to change their social media usage via financial incentives (Au, Ho & Chiu, 2021, p.3). However, many of these strategies have been largely ineffective at preventing the spread of particularly damaging forms of misinformation in the recent past. Specifically, it has been noted that during the 2020 election, a large amount of misinformation was spread across a variety of social media platforms regarding conspiracy theories about voting fraud (Chen, Chang, Lerman, Cowan & Ferrara, 2021, p.4). This misinformation was primarily spread widely on Twitter and Facebook, with a strong presence also on YouTube, and primarily consisted of claims that the election had been “stolen” in some way and that Donald Trump was the rightful winner of the election (Alba, 2021, p.1). Given that such serious misinformation has continued to spread despite a variety of attempts at counteracting it, it is clear that more research must be done in order to continue to disincentivize the spread of misinformation. One way to better analyze the complex system of incentives and existing power structures at play here is to use the framework of Actor-Network

Theory (ANT), as that allows us to make a map of the ways in which the different components of these structures interact (Crawford, 2020, p.1). ANT is a theoretical framework that labels all extant things in the social and natural worlds as existing in a constantly changing network of relationships, positing that all that exists are actors and the relationships between them. The following is an ANT graph that describes the preceding scenario:

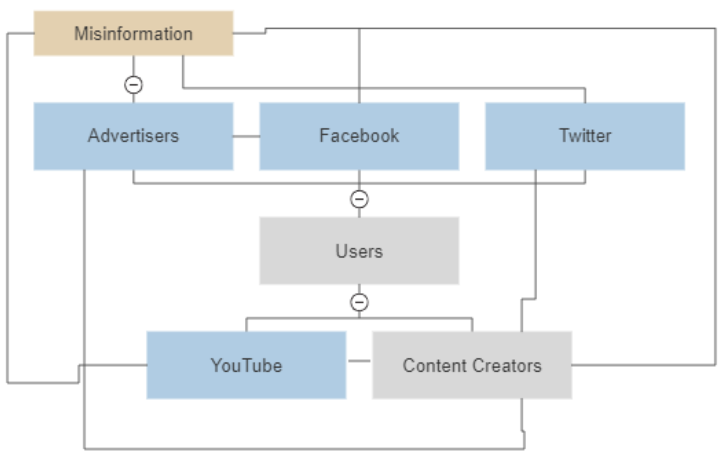

Figure 1: ANT Graph for Misinformation: Organizations, user groups, and concepts are labeled blue, grey, and orange respectively (Moreira, 2021)

As can be seen from Figure 1, it is clear that any incentive-based solution to reducing information spread needs to be based on either changing incentives for advertisers and platforms or for users and content creators, as those are the two key groups at play in this network (Valenzuela, Halpern, Katz, & Miranda, 2019, p.6). Given the previously described difficulties with platform-level solutions, it is clear that more research must be done to focus on ways to incentivize users of platforms to minimize posting misinformation. Using preliminary results from the technical report, it is possible that the best way to incentivize user behavioral change is to communicate to users that there will be potential rewards in their favor if they post authentic information that is well-sourced. This could manifest as a variety of different mechanisms, including direct financial incentives for posting well thought-out information or more indirect forms of persuasion including access to additional features on the website for being a member that is beneficial to the community. Having the incentive structure be focused on modifying user behavior also takes away the onus from large governments to actively regulate a vast and ever-changing social media landscape, instead allowing platforms themselves to incentivize the behavior independently and at a much faster rate than most governments can operate. The research for the STS project will thus necessitate the completion of some amount of the technical project in order to facilitate the explicit description of the user-based incentive solution that would be most beneficial, but this should not be difficult given that work has already begun on the technical project. The goal at the end of the STS project is to have a clear proposal for specific social media platforms such as Facebook or Twitter to modify their current misinformation content policies to incentivize users to post fact-checked and well-researched information rather than information that has been shared that solely exists to confirm users' existing biases in a variety of issues

GAME THEORY, PERSUASION, AND MISINFORMATION

In a variety of different contexts, the notion of persuasion is crucial to understanding the ways in which individual rational agents will interact with each other in order to maximize their own utility. The idea of persuasion describes any scenario in which the actions of one agent can send a signal to another agent to induce new actions in that other agent, and this framework has a variety of widespread potential applications in areas such as economics and politics. The technical project being proposed will serve to create a generalized algorithm to solve for the optimal strategy for a Sender in a Bayesian persuasion setting, in hopes that these results will generalize to other domains. In particular, the STS project that is tightly coupled to this notion of persuasion will analyze the ways in which social media platforms can modify their policies on misinformation to persuade users to verify the content of their posts to limit the amount of misinformation spread on said platforms. The eventual end goal of the STS project is to develop policy recommendations for specific social media platforms on specific misinformation policies. While the technical project has already begun, the results from the project will only allow for the proper development of the STS project by February 2022, when work on both projects will begin in earnest.

REFERENCES

Alba, D. (2021, October 14). *YouTube's stronger election misinformation policies had a spillover effect on Twitter and Facebook, researchers say.* The New York Times. https://www.nytimes.com/2021/10/14/technology/distortions-youtube-policies.html?searchResultPosition=2.

Arieli, I., & Babichenko, Y. (2019). Private bayesian persuasion. *Journal of Economic Theory*, *182*, 185–217. https://doi.org/10.1016/j.jet.2019.04.008

Au, C. H., Ho, K. K. W., & Chiu, D. K. W. (2021). Stopping healthcare misinformation: The effect of financial incentives and legislation. *Health Policy*, *125*(5), 627–633. https://doi.org/10.1016/j.healthpol.2021.02.010

Chen, E., Chang, H., Rao, A., Lerman, K., Cowan, G., & Ferrara, E. (2021). Covid-19 misinformation and the 2020 U.S. presidential election. *Harvard Kennedy School Misinformation Review*. https://doi.org/10.37016/mr-2020-57

Crawford, T. H. (2020). Actor-network theory. *Oxford Research Encyclopedia of Literature*. https://doi.org/10.1093/acrefore/9780190201098.013.965

Dughmi, S., & Xu, H. (2019). Algorithmic bayesian persuasion. *SIAM Journal on Computing*, *50*(3). https://doi.org/10.1137/16m1098334

Fallis, D. (2015). What is disinformation? *Library Trends*, *63*(3), 401–426. https://doi.org/10.1353/lib.2015.0014

Greenwald, A., Jafari, A., Gondek, D., & Ercal, G. (2001). *On no-regret learning, fictitious Play, and Nash equilibrium*. Brown University Department of Computer Science. Retrieved 2021, from http://cs.brown.edu/~amy/papers/icml.pdf

Icard, T. F. (2018). Bayes, bounds, and rational analysis. *Philosophy of Science*, *85*(1), 79–101. https://doi.org/10.1086/694837

Ito, J. (2011). Collective action for local commons management in rural Yunnan, China: Empirical evidence and hypotheses using evolutionary game theory. *Land Economics*, *88*(1), 181–200. https://doi.org/10.3368/le.88.1.181

Jasper, J. (2004). A strategic approach to collective action: Looking for agency in social-movement choices. *Mobilization: An International Quarterly*, *9*(1), 1–16. https://doi.org/10.17813/maiq.9.1.m112677546p63361

Kamenica, E., & Gentzkow, M. (2009). Bayesian persuasion. *American Economic Review*. https://doi.org/10.3386/w15540

Liao, X. (2021). Bayesian persuasion with optimal learning. *Journal of Mathematical Economics*. https://doi.org/10.1016/j.jmateco.2021.102534

McCloskey, D., & Klamer, A. (1995). One quarter of GDP is persuasion. *The American Economic Review*, *85*(2), 191–195.

Valenzuela, S., Halpern, D., Katz, J. E., & Miranda, J. P. (2019). The paradox of participation versus misinformation: Social media, political engagement, and the spread of misinformation. *Digital Journalism*, *7*(6), 802–823. https://doi.org/10.1080/21670811.2019.1623701

Wei, X., Yu, H., & Neely, M. J. (2020). Online primal-dual mirror descent under stochastic constraints. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, *4*(2), 1–36. https://doi.org/10.1145/3392157

West, D. M. (2017, December 18). *How to combat fake news and disinformation*. Brookings. Retrieved September 22, 2021, from https://www.brookings.edu/research/how-to-combat-fake-news-and-disinformation/.

Wu, L., Morstatter, F., Carley, K. M., & Liu, H. (2019). Misinformation in social media. *ACM SIGKDD Explorations Newsletter*, *21*(2), 80–90. https://doi.org/10.1145/3373464.3373475