An Empirical Analysis of Opinion Polarization Games Over Social Networks (ICML 2023)

Anonymous Authors[1]

Abstract

This paper provides a quantitative analysis of a game over a social network of agents, some of which are controlled by two players whose objectives are to maximize and minimize respectively polarization over this network. Opinions of agents evolve according to the Friedkin-Johnsen model, and players can change only the innate opinion of an agent of their choosing. Polarization is measured via the sample variance of the agents' steady-state opinions. The practically motivated constraint on the set of players' choice of agents to be disjoint transforms this simple zero-sum game into a compelling and largely unexplored research problem. We first analyze the functional properties of this game and characterize the optimal best response for each player given the agent. Next, we analyze the properties of the Nash equilibrium. Finally, we simulate a variation of the fictitious play algorithm to obtain the equilibrium in synthetic and real data networks, where the constraint mentioned above imposes a minor modification on the classical fictitious play algorithm. All of our codes and datasets are available publicly for research purposes.

1. Introduction

Opinion polarization has been one of the defining challenges of the current political era (Finkel et al., 2020; Iyengar et al., 2019; Allcott et al., 2020), yet the associated research problems on polarization dynamics over social networks remain largely unexplored. Several factors impact polarization, such as network size, structure/connectivity; agents' innate opinion distribution and behavioral/cognitive features; the presence of an exogenous influence, and many more. There are several important research questions with practical significance: For example, how to select a node to control in a network to minimize/maximize polarization in a competitive setting? It is in general challenging to investigate such questions in a purely theoretical framework. Hence, in this paper, we perform an empirical analysis of a constrained zero-sum game between two players, maximize and minimize polarization, respectively, by controlling the innate opinion of a chosen agent in a real social network. Our problem formulation builds on a few recent prior works (Chen & Racz, 2021) and (Musco et al., 2018) which respectively study maximizing and minimizing polarization by varying the innate opinions and network parameters in the presence of the Johnson-Friedkin opinion evolution model (Friedkin & Johnsen, 1990). To the best of our knowledge, this work is the first attempt to study the network polarization game over a social network.

There has been a significant amount of prior work on influence maximization over networks, e.g., (Richardson & Domingos, 2002; Kempe et al., 2003). This problem is identifying the most influential nodes in a network whose adoption of a product or an action will spread maximally in the social network. This line of work employs probabilistic propagation models, such as the independent cascade or the linear-threshold model, which specify how actions spread among social network individuals. At a high level, such propagation models classify the individuals in a network into active or inactive classes. They assume that active individuals influence their neighbors to become active according to specific probabilistic rules so that activity spreads in the network in a cascading manner. In this framework, it is natural to have more than one party who wants to spread specific information, e.g., for marketing or propaganda purposes. This imposes competition among players who aim to diffuse their preferred information on their product. Such competition can be modeled within a game-theoretic framework; hence, a characterization of the set of equilibria of such a game is sought after, e.g., (Dhamal et al.; Eshghi et al.).

In this paper, we are interested in the process of how individuals form opinions rather than how they adopt some actions. Opinion evolution resembles, rather than a discrete stochastic cascade as in such models, a dynamical process in which agents influence each other until convergence to a

[1]Anonymous Institution, Anonymous City, Anonymous Region, Anonymous Country. Correspondence to: Anonymous Author <anon.email@domain.com>.

Preliminary work. Under review by the International Conference on Machine Learning (ICML). Do not distribute.

steady-state.

A principal part of opinion dynamics is how nodes update their opinions over time. In the DeGroot model (DeGroot, 1974) each node updates its opinion using a weighted convex combination of its neighbors' opinions. The model developed by Friedkin and Johnsen (Friedkin & Johnsen, 1990) which is the model we base our work upon, considers that, in addition to its neighbors' opinions, a node also gives certain weightage to its innate opinion.

Posed in the competitive setting, opinion polarization game, offers a novel set of theoretical challenges. One of such

2. Preliminaries

2.1. Notation

Let $\mathcal{R}^n$ and $\mathcal{R}^{m\cdot n}$ denote the set of n-dimensional real vectors and the set of $m\cdot n$-dimensional real matrices, respectively. $\mathcal{N}$ represents the set of the positive integers, and $\mathcal{N}_0 = \mathcal{N}\cup\{0\}$. We define I as the identity matrix with proper dimension. We let $\mathbf{1}$ denote the vector of all ones. The superscript '$\top$' stands for the matrix transposition.

2.2. Model

We consider a social system with N agents, indexed by $i \in [N]$. The state/opinion of agent i at time t is denoted by $z_i(t) \in [0,1]$. The network is characterized by the triplet $(\mathcal{V}, \mathcal{E}, \mathcal{W})$ where $\mathcal{V}$, $\mathcal{E}$ and $\mathcal{W}$ represent the set of nodes (vertices), edges and weights. There are no assumptions imposed on the structure of the network.

Opinion Dynamics: We adopt the Johnson-Friedken opinion evolution model. Agent i has an innate opinion $s_i \in [0,1]$ and forms its opinion at time t, z_i, as a linear combination of that of network neighbors and its innate opinion as follows:

$$z_i(t) = \alpha_i s_i + (1-\alpha_i)\sum_{j\in V} w_{ij} z_j(t-1) \tag{1}$$

where $\alpha_i = 1/\left(1+\sum_{j\in V} w_{ij}\right)$ is a constant that ensures $z_i(t) \in [0,1]$. The limit (in time) of this process is referred in this work as steady-state[1] opinions, represented by $\tilde{z}_i = \lim_{t\to\infty} z_i(t)$. We let $\tilde{\mathbf{z}}$ to represent the vector of steady-state opinions, i.e., $\tilde{\mathbf{z}} = [\tilde{z}_1, \tilde{z}_2, \ldots, \tilde{z}_N]^T$. The following theorem[2] states that the steady-state opinions can be obtained from the vector of innate opinions via a linear transformation:

[1]We reserve the term "equilibrium" to denote the game-theoretic notion of Nash equilibrium.

[2]This result is well-known in the literature, see e.g. (Musco et al., 2018).

Theorem 2.1. *Given the dynamics in (1), $\tilde{\mathbf{z}} = \mathbf{Hs}$ where $\mathbf{H} = (\mathbf{I}+\mathbf{L})^{-1}$ is a doubly-stochastic matrix, and $\mathbf{L}$ is the Laplacian of the network.*

Polarization Metric: There are several polarization metrics used in the literature. In (Musco et al., 2018; Chen & Racz, 2021), two different metrics were employed to differentiate between "polarization" and "disagreement". Here, we adopt the network polarization metric in (Musco et al., 2018; Chen & Racz, 2021), which is simply the sample variance, defined as follows:

$$P(\tilde{\mathbf{z}}) \triangleq (\tilde{\mathbf{z}} - \overline{z})^T(\tilde{\mathbf{z}} - \overline{z}) \tag{2}$$

where $\overline{z} = \frac{1}{n}\sum_{v\in\mathbb{V}} \tilde{z}_v$ is the mean opinion averaged over the entire network. We note in passing that our results hold for a generic polarization function that is convex in $\tilde{\mathbf{z}}$.

Game Model: We consider a symmetric zero-sum game between two players, indexed by $m \in \{A, B\}$: player A (respectively, B), say Maximizer (resp., Minimizer), maximizes (resp., minimizes) $P(\tilde{\mathbf{z}})$ over its discrete-continuous hybrid action space $\mathcal{H}_A$ (resp., $\mathcal{H}_B$) which is defined as:

$$\mathcal{H}_m = \{(i^m, s_i^m) \mid s_i^m \in [0,1] \text{ for all } i^m \in \mathcal{S}_m\}, \tag{3}$$

where $\mathcal{S}_m = \{1, \cdots, N\} - \{i^{-m}\}$ is the discrete node set available to player m, s_i^m takes a value in $[0,1]$ interval.

The node index chosen by player m is denoted by $i^m \in [N]$, and the new value of the innate option is denoted by $s_i^m \in [0,1]$. Essentially, we consider the case where each players can modify one agent's innate opinion.

An important feature of this game arises from a practical consideration: Minimizer cannot choose the agent chosen by Maximizer, and vice versa. The action sets are constrained to be mutually exclusive at the agent level. In the pure strategies, this constraint corresponds to $i^m \neq i^{-m}$.

Such action set constrained zero-sum games, albeit its apparent importance in such practical security games, have not been explored in detail, except a few recent works that analyze the general setting where one players chosen action determines the available action set to the other player (Altman & Solan, 2009; Dhamal et al., 2018).

A distinction here to make, with respect to the common coupled constrained case in (Altman & Solan, 2009; Dhamal et al., 2018) is that we consider general strategy spaces, i.e., we allow mixed strategies while prior work on common coupled constrained games (Altman & Solan, 2009; Dhamal et al., 2018) formulate the game in pure strategies.

More formally, we define the strategy set of each player as $\sigma_m = [(i^m, p_{i^m}, s_i^m)]$ where p_{i^m} denotes the mixed strategy probability of the associated action (i^m, s_i^m) is chosen.

2.3. Fictitious Play

We utilize (discrete-time) fictitious play as a computational tool to determine the equilibria in our setting. Here, the players behave as if they are facing unknown, but stationary, distribution of opponents strategies. They form beliefs about opponent's strategy and play fully rationally with respect to these beliefs. We outline the learning process of fictitious belief for two players as follows:

Two players, indexed by $m = A, B$ are in a strategic form game with payoffs $J_m(a_m, a_{-m})$, at times $k = 1, 2, \ldots$. The number of times player m observes action J_{-m} before time k is denoted by $f_m^k(a_m, a_{-m})$, where $f_m^0(a_{-m})$ is a given initial condition. At time k, player m forms her belief on the opponent's stationary strategy as the empirical frequency of the past play (history), i.e.,

$$\mu_m^k(a_{-m}) = \frac{f_m^k(a_{-m})}{\sum\limits_{a'_{-m} \in \mathcal{A}_{-m}} f_m^k(a'_{-m})} \tag{4}$$

It is well understood that this procedure converges for several game classes including zero-sum, potential, 2×2 games and games solvable by iterated strict dominance, see e.g. Proposition 2.3 in (Fudenberg et al., 1998). Moreover, if it converges, it does so to the Nash equilibrium strategy of the associated strategy of the associated player. However, for a class of general-sum games which admit mixed strategies involving three or more actions, as shown in (), this procedure may not converge at all.

3. Main Results

Theorem 3.1. *For Maximizer, the optimal value, V_A, is either 0 or 1 regardless of the agent chosen i_A by Maximizer, and Maximizer's belief about Minimizer's strategy μ_A. For Minimizer, given his agent choice i_B, the optimal value is V_B.*

4. Algorithm

We simulate fictitious play between players A and B to obtain an equilibrium. At each round $k = 1, 2, \ldots$, players go through each node to find the optimal pure best response to the entire history of the opponent, as described in the previous section. In the following, we present the optimal values, i.e., V_m for a given selected agent i_m and each player $m = A, B$.

With 0 and 1 two options for n nodes in the network, in total Maximizer have $2n$ options. Maximizer will go through each node and find optimal decision (i, v_A) that maximize polarization by checking all $2n$ possibilities. In (Chen & Racz, 2021), it proved that the adversary's optimization problem is a convex maximization problem in s_i, therefore, the value (v_A) Maximizer used to replace innate opinion(s_i) of chosen agent must be extreme: v_A = 0 or 1. Also, the optimal value for Minimizer to minimize polarization by responding to each Maximizer's action can be derived as:

$$s_i^* = \frac{-\sum\limits_{j \neq i} s_j (a_j - c)^T (a_i - c)}{(a_i - c)^T (a_i - c)} \tag{5}$$

To play fully rationally with respect to Maximizer's past history, the optimal value V_B for Minimizer is the weighted average of all s_i^*. i.e.,

$$V_B = \sum_{l=0}^{2n} s_i^* \mu_{Al}^k (a_{-Al}) \tag{6}$$

In the end of the iterations, player will find optimal decision (i, v_m) that minimize/maximize the steady-state polarization.

Definition 4.1. For both Maximizer and Minimizer, the optimal action at round k is the one that maximize/minimize expected payoff $\mathbb{E}(J_m)^k$.

At time k, player m will take action to maximize/minimize the steady-state polarization J_m by responding to it's opponent's action a_{-m}. If the opponent takes mixed strategy, denoting opponent's action as a_{-ml}, then player m will maximize/minimize the expected payoff $\mathbb{E}(J_m)$ according to her belief of her opponent's strategy. Recall from Definition **??** that the beliefs are updated after each round of the game. i.e., At each round k, the expected payoff is:

$$\mathbb{E}(J_m)^k = \sum_{l \in V} \mu_{-ml}^k J_{ml} \tag{7}$$

Player m then chooses the action a_m that maximizes her expected payoff $\mathbb{E}(J_m)^k$ replacing a_{-m} with her belief $\mu_m^k(a_{-m})$, i.e.,

$$a_m = \arg\max_{a_m \in \mathcal{A}_m} \mathbb{E}(J_m, \mu_m^k) \tag{8}$$

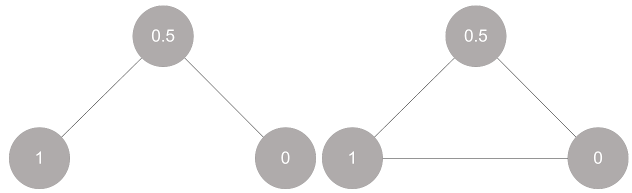
5. Numerical Results

This section we present our running example and simulation configurations. Some assumptions will be adapted throughout all the experiments. The simulation result motivate our theoretical result of Nash equilibrium proof – (the unique mixed NE). We use both synthetic and real data in this paper. All datasets and codes are available at: `https://anonymous.4open.science/r/Network_Node_Selection-BCC2/README.md`.

Algorithm 1 Fictitious Play Algorithm

Input: innate opinion vector (s)
Hyper-parameter: k, r
Output:($G_A^k, G_B^k, G_A^{-r}, G_B^{-r}$)

Take player m's random action (i_m, v_m)
Update innate opinion vector(s) to $s_m \leftarrow$ replace ith element in opinion vector s to v_m
Compute steady state polarization $\leftarrow P(z_m) = f(s_m)$ (see **Eq.(2)**)
Update the empirical frequency of the player m's play G_m^0(see **Eq.(4)**), $(m = A, B)$
Given k = 0, Game rounds: $\#$ of game the algorithm runs
while $k \leq$ Game rounds **do**
  if G_m^k (m = A,B) do not change anymore **then**
    Break
    Output (G_A^k, G_B^k)
  else
    $k = k + 1$
    Maximizer take action(i, v_A) that maximize steady-state polarization by responding to Minimizer's action history(G_B^{k-1})
    Update the empirical frequency of the Maximizer's past play (G_A^k) (see **Eq. 8**)
    Minimizer take action(j, v_B) that minimize steady-state polarization by responding to Maximizer's action distribution(G_A^k)
    Update the empirical frequency of the Minimizer's past play(G_B^k)
  end if
end while
return output



(a) Network 1 (b) Network 2

Figure 1. Synthetic Networks

Data	Size	Network	Polariz.	Source
Karate club	34	undirected	0.16	(Zachary, 1977)
Reddit	553	undirected	0.01	(De et al., 2014)
Twitter	548	undirected	0.17	(De et al., 2014)

Table 1. Details of datasets used in our simulations.

5.1. Data Description

5.1.1. Synthetic Data

The extreme network includes three agents. Since all opinions are in the range of $[0, 1]$, we took two extreme opinions and a middle opinion. Three agents innate opinion are set to

$$\mathbf{s} = [1, 0.5, 0]^T \tag{9}$$

In addition, we set all the edge weights in the network equal which leads to (due to (1))

$$w_{ij} \begin{cases} = 0.02, & \text{if } (\mathrm{v_i}, \mathrm{v_j}) \in \mathbb{E} \\ = 0, & \text{otherwise;} \end{cases}$$

5.1.2. Real Data

We use three real social networks as shown in Table 1. The Karate club is a network of 34 members. The Reddit[3] and Twitter datasets are online social network snapshots of 556 and 548 users, respectively. Three isolated vertices are discarded from the Reddit network since they do not yield any consequences for the opinion formation process, yielding 553 nodes. For all three networks, we obtain the adjacency matrix from the datasets. The innate opinion for Reddit and Twitter are also provided in (De et al., 2014). We generate one innate opinion vector $\mathbf{s}_{karate}$ for Karate club network with each entry randomly drawn from range [0,1], similar to the experiment in (Biondi et al., 2022).

The Karate network is less clustered compared to the other two networks. The Reddit and Twitter networks are bigger and more clustered. The Reddit network has a big

[3]Original Reddit data includes 556 nodes, 3 of them are isolated. The isolated nodes can affect the polarization since they are not influenced by the opinion dynamic, we removed these three nodes in our experiment.

	Maximizer				Minimizer				
	i^A	s_i	s_i^A	p_{i^A}	i^B	s_i	s_i^B	p_{i^B}	$E(J)$
NE 1	1	0.5	0	1	2	0	0.5	1	0.481
NE 2	1	0.5	1	1	0	1	0.5	1	0.273
NE 3	1	0.5	0	0.48	0	1	0.26	0.52	0.273
	1	0.5	1	0.52	2	0.5	0.76	0.48	

Table 2. Network 1 simulation result.

	Maximizer				Minimizer				
	i^A	s_i	s_i^A	P	i^B	s_i	s_i^B	P	$E(J)$
NE 1	1	0.5	1	1	0	1	0.5	1	0.453
NE 2	0	1	1	1	1	0.5	0.5	1	0.445
NE 3	1	0.5	0	5	2	0	0.5	1	0.445
NE 4	1	0.5	0	0.5	2	0	0.25	0.5	0.26
	1	0.5	1	0.5	2	0	0.75	0.5	

Table 3. Network 2 simulation result.

community, and the Twitter network has four well-depicted communities.

5.2. Numerical Results

5.2.1. Synthetic Network

We start the game with different random actions from both Maximizer and Minimizer. At the first round, players can choose any agent in the network and randomly change it's opinion to 0 or 1. We call players' random actions at first round 'initial condition' of the game. After first-round game, players started fictitious play with considering opponent's history from the last round. The constraint are 1) two players cannot choose the same agent in the same round, 2) players cannot choose the agent that opponent touched in the last round. We set innate opinions s (see (9)) and experimented with random initial conditions. After running the simulation that illustrated in Algorithm 1, as the Table 2 shows, the fictitious game give us an unique mixed Nash equilibrium (NE) for each network regardless of the initial conditions.

Example 1. We start game on a three-nodes synthetic network where one node is connected to the other two agents and weight equal to 0.02, as shown in Figure **??**. For the innate opinion equal to S (innate opinion is marked on the node), one unique mixed Nash equilibrium exist.

As shown in Table 2, network 1 has two pure Nash equilibrium and one mixed Nash equilibrium. At NE_1, the network polarization is 0.481 ; At NE_2, the network polarization is 0.453; At the mixed NE_3, if $m = A$, given $\mu_{A1} = 0.49$ and $\mu_{A2} = (1 - \mu_{A1}) = 0.51$, we have expected payoff 0.273 for both Maximizer's actions. If $m = B$, given $\mu_{B1} = 0.52$ and $\mu_{B2} = (1 - \mu_{B1}) = 0.48$, we have expected payoff 0.273 for both Minimizer's actions.

Example 2. We start game on a three-nodes synthetic network where three nodes are connected to the each other two with a weight equal to 0.02, as shown in Figure **??**. For the innate opinion equal to S (innate opinion is marked on the node), one unique mixed Nash equilibrium exist.

As shown in Table 3, network 2 has three pure Nash equilibrium and one mixed Nash equilibrium. At pure NE_1, the network polarization is 0.453; At NE_2, the network polarization is 0.445; At pure NE_3, the network polarization is 0.445. At the NE_4, if $m = A$, given $\mu_{A1} = 0.5$ and $\mu_{A2} = (1 - \mu_{A1}) = 0.5$, we have expected payoff 0.26 for both Maximizer's actions. If also use $m = B$, Minimizer empirical frequency has given $\mu_{B1} = 0.5$ and $\mu_{B2} = 0.5$, we have the expected payoff equal to 0.26 for both actions.

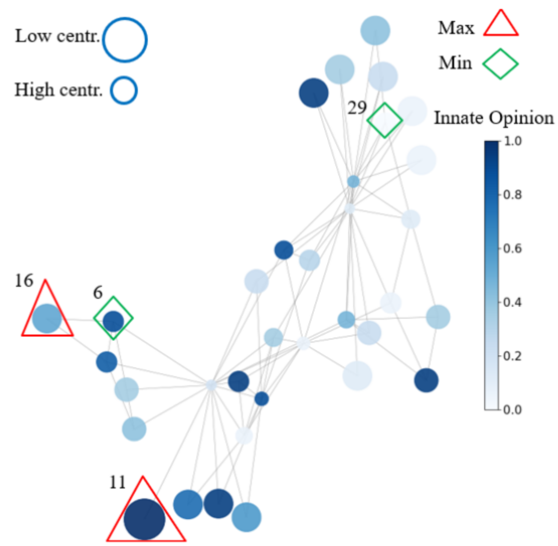


Figure 2. Karate Network mixed NE. The node color intensity represents the node's opinion value in the range [0,1]. Larger circle means this node has less connections with other nodes. At equilibrium, Maximizer selected nodes that are circled with triangle shape; Minimizer selected nodes that are circled with square shape.

5.2.2. Real Network

This section present results of polarization game on Karate, Reddit, and Twitter network. We start the game on with random initial action from both Maximizer and Minimizer. To analyze the maximizer and minimizer's behavior tendency of selecting NE, we analyze selected nodes with three network centrality concepts: Degree Centrality ($C1$), Eigen Value Centrality($C2$), Closeness Centrality($C3$).

a)**Karate Network.** In the Karate network, one unique mixed Nash equilibrium exist. Maximizer takes action (agent, opinion), and Minimizer takes action (agent, opinion) when the game reaches to Nash equilibrium, as show in simulation result Table 4.

	Maximizer				Minimizer				
	i^A	s_i	s_i^A	P	i^B	s_i	s_i^B	P	E(J)
NE 1	11	0.77	1	0.75	6	0.72	0	0.9	0.06
	16	0.51	1	0.25	29	0.21	1	0.1	

Table 4. Karate Network simulation result.

.

	Maximizer		Minimizer	
Nodes	11	16	6	29
C1 rank	1	6	20	24
C2 rank	2	1	9	18
C3 rank	3	1	17	21

Table 5. Centrality analysis for Karate network NE

As shown in Table 4, if $m = A$, given $\mu_{A1} = 0.75$ and $\mu_{A2} = (1 - \mu_{A1}) = 0.25$, we have expected payoff 0.059 for both Maximizer's actions. If also use $m = B$, Minimizer's empirical frequency provided $\mu_{B1} = 0.25$ and $\mu_{B2} = (1 - \mu_{B1}) = 0.75$, we have expected payoff 0.059 for both actions.

b) **Reddit Network** The experiments' result showed two different Nash equilibrium in the Reddit Network. One pure Nash equilibrium and one mixed Nash equilibrium.

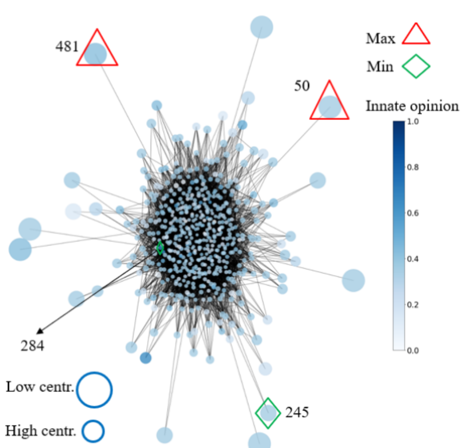


Figure 3. Reddit Network Mixed NE.

As shown in Table 6, if m = A, given $\mu_{A1} = 0.29$ and $\mu_{A2} = (1 - \mu_{A1}) = 0.71$, we have expected payoff 0.068 for both Maximizer's actions. If also use $m = B$, given $\mu_{B1} = 0.84$ and $\mu_{B2} = (1 - \mu_{B1}) = 0.16$, we have expected payoff 0.068 for both Minimizer's actions.

c) **Twitter** The experiments find one mixed Nash equilibrium where Maximizer takes action (agent, opinion) and Minimizer takes action (agent, opinion). As shown in Table 8.

	Maximizer				Minimizer				
	i^A	s_i	s_i^A	P	i^B	s_i	s_i^B	P	E(J)
NE 1	50	0.5	1	0.29	245	0.41	0.5	0.84	0.07
	481	0.5	1	0.71	284	0.63	0.29	0.16	

Table 6. Reddit Network simulation result.

	Maximizer		Minimizer	
Nodes	50	481	245	284
C1 rank	1	6	10	>50
C2 rank	5	7	13	>50
C3 rank	5	7	13	>50

Table 7. Centrality analysis for Reddit network NE

	Maximizer				Minimizer				
	i^A	s_i	s_i^A	P	i^B	s_i	s_i^B	P	E(J)
NE 1	202	0.56	0	0.45	529	0.49	0.76	0.4	0.28
	351	0.64	0	0.45	37	0.61	0.8	0.4	
	490	0.57	0	0.1	199	0.23	0.7	0.2	

Table 8. Twitter Network simulation result.

	Maximizer			Minimizer		
Nodes	202	351	490	529	37	199
C1 rank	8	11	16	>50	>50	>50
C2 rank	1	8	25	>50	>50	>50
C3 rank	1	25	2	>50	>50	>50

Table 9. Centrality analysis for Twitter network NE

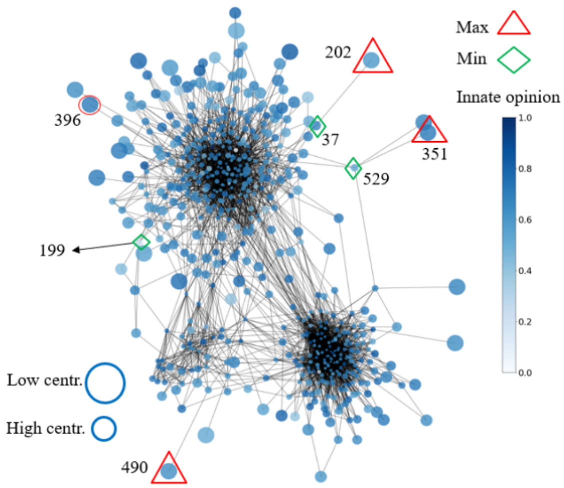


Figure 4. Mixed NE on Twitter Network

As shown in Table 8, if m = A, given the probability of three actions: $\mu_{A1} = 0.4$, $\mu_{A2} = 0.4$, and $(1-\mu_{B1}-\mu_{B2}) = 0.2$, we have expected payoff 0.282 for both Maximizer's actions. If also use $m = B$, given the probability of three actions: $\mu_{B1} = 0.45$, $\mu_{B2} = 0.45$, and $(1 - \mu_{B1} - \mu_{B2}) = 0.1$, we have expected payoff 0.282 for all Minimizer's actions.

5.2.3. MAXMIN RESULT

This section presents the result of polarization game with pure maxmin strategy, where maximizer choose its best action(node, opinion) while knowing what action minimizer will take after its action. As mentioned in section 4, (Chen & Racz, 2021) proved that the adversary's optimization problem is a convex maximization problem in s_i. Hence for each agent, maximizer will choose to change its opinion to 0 or 1. In total, maximizer have $2n$ possible actions in a n-agent network. For each maximizer's possible action, Minimizer will go through all available agents(excludes the one maximizer has chosen) and take best action to minimize the network polarization. As mentioned in eq. 18, for a chosen agent, minimizer will update its opinion to s^*. The polarization will be calculated with new opinion array(s') after both players' action as show in Eq.2. Then maximizer will choose the action that gives maximum polarization among all $2n$ polarizations. Table 10-14 gives the numeric result of maxmin strategy.

Table 10 shows the maxmin strategy on Karate network. Maximizer will choose agent 16 and change its innate opinion 0.51 to 1 by knowing minimizer' action, at the same time, minimizer choose agent 6 and change its innate opinion 0.72 to 0. The polarization after maxmin strategy is 0.17, which has increased than karate network's original polarization 0.16. In the Reddit network, as Table 12 shown, maximizer will choose agent 481 and update its innate opinion from 0.5 to 1 while minimizer choose agent 284 and update its innate opinion 0.63 to 0. The polarization of Reddit network was 0.01 before the player's action, it increases to 0.07 after maxmin strategy. Table 14 shows maxmin strategy on Twitter newtwork. Maximizer chooses agent 481 and update its innate opinion 0.5 to 1 while minimizer select agent 245 and change its innate opinon 0.41 to 0.5. The polarization after both players' action is 0.07, which has decreased compared to 0.17, the original polarization of Twitter network.

	Maximizer	**Minimizer**
Nodes	16	6
Innate opinion	0.51	0.72
New opinion	1	0
Polarization	0.17	

Table 10. Karate Network Maxmin result.
.

	Maximizer	**Minimizer**
Nodes	16	6
C1 rank	5	19
C2 rank	0	16
C3 rank	0	8
rank(Op - mean)	4	26

Table 11. Centrality analysis for Karate network Maxmin

5.2.4. MINMAX RESULT

This section illustrates the outcome of a polarization game with a pure minimax strategy, where the minimizer chooses the most effective action (node, opinion) while anticipating the next move the maximizer will make. The convex maximization issue in the fictitious play and Maxmin strategy makes it possible to derive the minimizer's optimal value after the maximizer acts. By providing the minimizer a set of 11 discrete actions, we can streamline the game and operation the minimizer to act first. Alternative V_B values for the minimizer include 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9,

	Maximizer	**Minimizer**
Nodes	481	284
Innate opinion	0.5	0.63
New opinion	1	0
Polarization	0.07	

Table 12. Reddit Network Maxmin result.
.

	Maximizer	Minimizer
Nodes	481	284
C1 rank	5	>¿50
C2 rank	6	>¿50
C3 rank	6	>¿50
rank(Op - mean)	256	543

Table 13. Centrality analysis for Reddit network Maxmin

	Maximizer	Minimizer
Nodes	396	294
Innate opinion	0.65	0.67
New opinion	0	1
Polarization	0.26	

Table 14. Twitter Network Maxmin result.

.

	Maximizer	Minimizer
Nodes	396	294
C1 rank	11	>50
C2 rank	2	>50
C3 rank	4	>50
New Opinion	284	398

Table 15. Centrality analysis for Twitter network Maxmin

	Maximizer	Minimizer
Nodes	11	6
Innate opinion	0.77	0.72
New opinion	1	0.1
Polarization	0.18	

Table 16. Karate Network Minmax result.

.

	Maximizer	Minimizer
Nodes	11	6
C1 rank	0	19
C2 rank	2	16
C3 rank	1	8
rank(Op - mean)	33	26

Table 17. Centrality analysis for Karate network Minmax

and 1. In an n-agent network, the minimizer has $11n$ total potential actions. The network's best course of action will be decided for each minimizer's potential action according to the maximizer after reviewing all available agents (apart from the one the minimizer has selected). Table 16-20 gives the numeric result of minmax strategy.

Table 16 shows the minimax strategy on the Karate network. Minimizer will choose agent 11 and change its innate opinion from 0.77 to 1 by knowing the minimizer's action, at the same time, minimizer chooses agent 6 and changes its innate opinion from 0.72 to 0.1. The polarization after minimax strategy is 0.18, which has increased from than karate network's original polarization of 0.16. In the Reddit network, as Table 12 shows, the maximizer will choose agent 481 and update its innate opinion from 0.5 to 1 while the minimizer chooses agent 245 and updates its innate opinion from 0.41 to 0.5. The polarization of Reddit network was 0.01 before the player's action, it increases to 0.07 after minmax strategy[4]. Table 14 shows the minimax strategy on the Twitter network. Maximizer chooses agent 351 and updates its innate opinion from 0.64 to 0 while the minimizer selects agent 378 and changes its innate opinion from 0.77 to 0. The polarization after both players' actions is 0.28, which has increased compared to 0.17, the original polarization of the Twitter network.

6. Discussion

In this section, we present the results of our study on the interplay between network structure and the evolution of

[4]In Reddit network, the polarization after the minimax strategy is a round-up of the actual value 0.068. The polarization after the maximin strategy is a round-up of the actual value 0.067.

	Maximizer	Minimizer
Nodes	481	245
Innate opinion	0.5	0.41
New opinion	1	0.5
Polarization	0.07	

Table 18. Reddit Network Minmax result.

.

	Maximizer	Minimizer
Nodes	481	245
C1 rank	5	9
C2 rank	6	12
C3 rank	6	12
rank(Op - mean)	256	525

Table 19. Centrality analysis for Reddit network Minmax

	Maximizer	Minimizer
Nodes	351	378
Innate opinion	0.64	0.77
New opinion	0	0
Polarization	0.28	

Table 20. Twitter Network Minmax result.

.

	Maximizer	Minimizer
Nodes	351	378
C1 rank	10	>50
C2 rank	7	>50
C3 rank	6	>50
rank(Op - mean)	212	525

Table 21. Centrality analysis for Twitter network Minmax

opinion polarization in social networks. Specifically, we analyze the Nash equilibrium, pure strategy maxmin, and pure strategy minimax of three different networks. We then examine the centrality and innate opinion of selected nodes and present the results in terms of centrality rank and opinion distance rank. Centrality rank is determined by three metrics, including Degree Centrality ($C1$), Eigen Value Centrality($C2$), and Closeness Centrality($C3$). Rank 1 represents the node with the lowest centrality among all nodes in the network. Opinion distance is the difference between the innate opinion of selected nodes and the mean opinion of the network. Rank 1 in opinion distance represents the node with the innate opinion that is closest to the mean opinion of the network. Both ranks are ascending, starting with smaller centrality and smaller opinion gaps.

Tables 4-21 display the Nash equilibrium result. The centrality analysis tables, including Table **??**, Table **??**, and Table **??**, are presented with the Nash Equilibrium. In both Karate and Reddit networks, the maximizer chooses nodes from the lowest 7 centralities, while nodes selected by the minimizer have relatively high centrality. This trend is more pronounced in the Twitter NE result, where all three nodes in the minimizer's NE mixed strategy have high centrality rank larger than 50, while the maximizer chooses nodes from the lowest 25 centralities. Similar behavior tendencies are also shown in the pure strategy maximin and minimax results.

For instance, Table 10-20 show all strategies found with the Maxmin and Minmax method. In the Karate network, the maximizer selects node 11 with the property that all three centralities are ranked in the lowest 3 ranks, while the minimizer selects node 6, whose innate opinion is far away from the mean opinion (with rank 26 out of 33). It is more evident in the Twitter and Reddit networks that the maximizer chooses nodes with low centrality, while the minimizer chooses nodes with centrality ranked higher than 50. Furthermore, these nodes all have an innate opinion that is far away from the mean opinion of the network they belong to.

We offer an intuitive explanation for these results. The F-J model is an averaging dynamics, which means that it may mitigate the polarization influence of well-connected agents' innate opinions. Therefore, the maximizer may select relatively isolated agents to amplify polarization across the network. As a result, the averaging dynamic may have a smaller impact on the maximizer's choice of agents. For the minimizer, since the overall averaging dynamic is helping to reduce the network polarization, additional factors beyond the network topology, such as the extremity of the original innate opinion, may influence their agent selection strategy, which is supported by our result that the innate opinions of minimizer selected nodes are relatively extreme compared to other nodes in the network. However, the extent to which these factors influence the minimizer's choice may depend on the characteristics of the network under consideration.

7. Conclusion

In this paper, we have provided a quantitative analysis of a game played over a social network of agents, where the goal of two players is to maximize and minimize polarization over the network. We have utilized the F-J model to simulate the evolution of the opinions of the agents and allowed the players to change only the innate opinion of an agent of their choosing. The game is subject to a practically motivated constraint, limiting the set of players' choice of agents to be disjoint, which creates a unique and complex problem that has not been extensively explored in the literature.

Our study characterizes the functional properties of this game, including the optimal best response for each player given the agent, the properties of the Nash equilibrium, and a novel variation of the fictitious play algorithm that takes into account the practical constraint of disjoint agent selection. Our simulation results on synthetic and real data networks demonstrate the novelty of our approach and provide useful insights into the behavior of the game under different scenarios.

The contribution of this work lies in the development of a novel approach to analyzing games over social networks, which can have important implications for real-world applications such as political campaigns, social movements, and marketing strategies. Moreover, the numerical results reveal the behavior tendencies of the two players, where the maximizer tends to choose nodes with lower centrality, while the minimizer prefers nodes with extreme opinions. These insights are particularly valuable given the computationally expensive nature of simulating this game, as they can help to develop more efficient heuristic algorithms that require less computing power.

References

Allcott, H., Boxell, L., Conway, J., Gentzkow, M., Thaler, M., and Yang, D. Polarization and public health: Partisan differences in social distancing during the coronavirus pandemic. *Journal of public economics*, 191:104254, 2020.

Altman, E. and Solan, E. Constrained games: The impact of the attitude to adversary's constraints. *IEEE Transactions on Automatic Control*, 54(10):2435–2440, 2009.

Biondi, E., Boldrini, C., Passarella, A., and Conti, M. Dynamics of opinion polarization. *arXiv preprint arXiv:2206.06134*, 2022.

Chen, M. F. and Racz, M. Z. An adversarial model of network disruption: Maximizing disagreement and polarization in social networks. *IEEE Transactions on Network Science and Engineering*, 2021.

De, A., Bhattacharya, S., Bhattacharya, P., Ganguly, N., and Chakrabarti, S. Learning a linear influence model from transient opinion dynamics. In *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*, pp. 401–410, 2014.

DeGroot, M. H. Reaching a consensus. *Journal of the American Statistical association*, 69(345):118–121, 1974.

Dhamal, S., Ben-Ameur, W., Chahed, T., and Altman, E. Optimal investment strategies for competing camps in a social network: A broad framework. *IEEE Transactions on Network Science and Engineering, DOI: 10.1109/TNSE.2018.2864575.*

Dhamal, S., Ben-Ameur, W., Chahed, T., and Altman, E. Resource allocation polytope games: Uniqueness of equilibrium, price of stability, and price of anarchy. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.

Eshghi, S., Preciado, V., Sarkar, S., Venkatesh, S., Zhao, Q., D'Souza, R., and Swami, A. Spread, then target, and advertise in waves: Optimal budget allocation across advertising channels. *IEEE Transactions on Network Science and Engineering, DOI: 10.1109/TNSE.2018.2873281.*

Finkel, E. J., Bail, C. A., Cikara, M., Ditto, P. H., Iyengar, S., Klar, S., Mason, L., McGrath, M. C., Nyhan, B., Rand, D. G., et al. Political sectarianism in america. *Science*, 370(6516):533–536, 2020.

Friedkin, N. E. and Johnsen, E. C. Social influence and opinions. *Journal of Mathematical Sociology*, 15(3-4): 193–206, 1990.

Fudenberg, D., Drew, F., Levine, D. K., and Levine, D. K. *The theory of learning in games*, volume 2. MIT press, 1998.

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., and Westwood, S. J. The origins and consequences of affective polarization in the united states. *Annual Review of Political Science*, 22(1):129–146, 2019.

Kempe, D., Kleinberg, J., and Tardos, É. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 137–146, 2003.

Musco, C., Musco, C., and Tsourakakis, C. Minimizing polarization and disagreement in social networks. Proceedings of the World Wide Web Conference, WWW 2018, pp. 369–378, April 2018. doi: 10.1145/3178876.3186103.

Richardson, M. and Domingos, P. Mining knowledge-sharing sites for viral marketing. In *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 61–70, 2002.

Zachary, W. W. An information flow model for conflict and fission in small groups. *Journal of anthropological research*, 33(4):452–473, 1977.

A. Proof of Theorem 1

The J-F dynamics lead to

$$z = Hs \tag{10}$$

as a steady-state point where $H = (I + L)^{-1}$ is a doubly-stochastic matrix.

Plugging (10) into (2), we obtain polarization as a function of innate opinions:

$$J(s) = (Hs - c^T Hs)^T (Hs - c^T Hs) \tag{11}$$

where we define $\mathbf{c}_n \triangleq \frac{1}{n}\mathbf{1}_n$ and $\mathbf{1}_n$ refers to the $n \cdot 1$ 1s vector.

We next note that $\mathbf{c}^T Hs = \mathbf{c}^T s$, due to H being column-stochastic, i.e., $c^T H = c^T$, hence we have

$$J(s) = (Hs - c^T s)^T (Hs - c^T s) \tag{12}$$

Towards computing the derivative $\frac{dJ}{ds}$, we have

$$J(s + \epsilon e_i) - J(s) = \tag{13}$$

$$\sum_k p_k (Hs^k - c^T s^k + \epsilon(He_i - c^T e_i))^T (Hs^k - c^T s^k + \epsilon(He_i - c^T e_i)) - (Hs^k - c^T s^k)^T (Hs^k - c^T s^k) = \tag{14}$$

$$\epsilon^2 (He_i - c^T e_i)^T (He_i - c^T e_i) + \epsilon((Hs - c^T s^k)^T)(He_i - c^T e_i) + (He_i - c^T e_i)^T (Hs^k - c^T s^k)) \tag{15}$$

We next compute

$$\begin{aligned} \frac{dJ}{ds} &= \lim_{\epsilon \to 0} \frac{J(s + \epsilon e_i) - J(s)}{\epsilon} \\ &= 2(Hs - c^T s)^T (He_i - c^T e_i) \end{aligned} \tag{16}$$

.

Noting that $He_i = h_i$ and $c^T e_i = \frac{1}{n}$, we obtain the locally optimal point s_i^* as the solution of

$$\left(s_i h_i - s_i c + \sum_{j \neq i} (s_j h_j - c_j s_j) \right)^T (h_i - c) = 0 \tag{17}$$

Using the column-stochasticity of H, we obtain s_i^* as

$$s_i^* = \frac{- \sum_{j \neq i} s_j (h_j - c)^T (h_i - c)}{(h_i - c)^T (h_i - c)} \tag{18}$$