General Opinion Formation Games with Social Group Membership (Short Paper)*

Vittorio Bilò[1,†], Diodato Ferraioli[2,†] and Cosimo Vinci[2,*,†]

[1]Università del Salento, Italy
[2]Università degli Studi di Salerno, Italy

Abstract

Modeling how agents form their opinions is of paramount importance for designing marketing and electoral campaigns. In this work, we present a new framework for opinion formation which generalizes the well-known Friedkin-Johnsen model by incorporating three important features: (i) social group membership, that limits the amount of influence that people not belonging to the same group may lead on a given agent; (ii) both attraction among friends, and repulsion among enemies; (iii) different strengths of influence lead from different people on a given agent, even if the social relationships among them are the same.

We show that, despite its generality, our model always admits a pure Nash equilibrium which, under opportune mild conditions, is even unique. Next, we analyze the performances of these equilibria with respect to a social objective function defined as a convex combination, parametrized by a value $\lambda \in [0, 1]$, of the costs yielded by the untruthfulness of the declared opinions and the total cost of social pressure. We prove bounds on both the price of anarchy and the price of stability which show that, for not-too-extreme values of λ, performance at equilibrium are very close to optimal ones. For instance, in several interesting scenarios, the prices of anarchy and stability are both equal to $\frac{\max\{2\lambda,1-\lambda\}}{\min\{2\lambda,1-\lambda\}}$ which never exceeds 2 for $\lambda \in [1/5, 1/2]$.

Keywords

Opinion Formation Games, Pure Nash Equilibria, Price of Anarchy, Price of Stability

1. Introduction

In recent years, a lot of attention has been devoted to studying how people form their opinions, and how the social media affect the opinion formation process. Understanding these aspects is of fundamental importance for analysing and forecasting electoral flows and implement suitable electoral campaigns, or for marketing purposes.

Most of the approaches proposed in the literature usually assume that people try to "imitate" their "friends". This is, for example, the case of the celebrated DeGroot (DG) model [2, 3],

Proceedings of the 23rd Italian Conference on Theoretical Computer Science, Rome, Italy, September 7-9, 2022

*This work was partially supported by the Italian MIUR PRIN 2017 Project ALGADIMAR "Algorithms, Games, and Digital Markets" and by "GNCS–INdAM".

An extended version of this work appears in the proceedings of IJCAI 2022 [1].

*Corresponding author.

†These authors contributed equally.

vittorio.bilo@unisalento.it (V. Bilò); dferraioli@unisa.it (D. Ferraioli); cvinci@unisa.it (C. Vinci)

where opinions are continuous and repeatedly updated to the average of the opinions expressed by one's friends. Among the most relevant generalizations of the DG model is the one of Friedkin-Johnsen (FJ) [4], in which people have an internal belief about the matter in object that limits in some way the influence of friends. Other approaches consider discrete opinion spaces [5, 6], or limited/local interactions [7, 8], or dynamic settings where social relationships and internal beliefs evolve over time [9, 10, 11, 12, 13].

All these models, however, focus on imitative behaviour only. Indeed, there are many examples in which our opinion is not only influenced by imitation of our friends, but also by rejection of our "enemies". One example arises from youth subcultures, where peoples belonging to two different subcultures, even if a strict relation exists among them (e.g., they are relatives or they are in the same school), try to make opposite choices about style and interests, with the goal to distinguish each from the other. Another example comes from politics, where the position of a party about a topic sometimes arises more in opposition to adversaries rather than from principles and values. To the best of our knowledge, very few works considered this mixture of attraction and repulsion in opinion formation [14, 15] and, in any case, they limit the modelling of attraction/repulsion to a logic setting, which can only be applied to discrete opinions.

Both examples described above also highlight a fundamental feature of opinion formation that most of the discussed works neglect: membership in social groups. Indeed, followers of a subculture (e.g., hipsters) are used to limit their musical interests to the genre of reference of this subculture (e.g. indie), even if they are influenced by people listening to different music styles. Similarly, people belonging to a party usually support only opinions "allowed" by that party, despite the amount of social pressure they may face.

Yet another limitation of most of the considered models is that they assume a strength of attraction (or dis-attraction) that is the same for each pair of friends (enemies), possibly diversified only by a scaling factor measuring the weight of the social relationship. However, it may not be the case that hipster guys are attracted in the same way by emo peers and by geek peers, even if they all share the same social relationship. Similarly, the position of a right party on a given topic may be influenced in different ways by a center party or by an extreme-right party, even if the right party shares the same contacts with the other two (e.g., they are always allied at elections). Such degree of generality, although only restricted to attraction phenomenon, has been considered before only by [10].

1.1. A New Model of Opinion Formation

In this work, we tackle all the above limitations by proposing a new, general, model in which people choose their opinion by trying to simultaneously imitate their "friends" and distinguish themselves from their "enemies". We allow opinions to be chosen from a continuous set (differently from [14, 15]), and model social group membership by limiting the set of choices of each agent within the boundaries imposed by her social group. Finally, we also allow the strength of attraction and repulsion to be completely arbitrary and pair-specific, and not only influenced by the weights of the social relationships.

Specifically, we model this opinion formation framework as a cost minimization game with n agents, in which each agent belonging to a social group chooses an opinion whose distance

from her private belief cannot exceed a certain threshold yielded by the boundaries of the group. In other words, while an agent is allowed to change her/his opinion, this opinion cannot lead this agent too far away from the cluster (social group) she/he feels to belong to. We model these constraints by imposing that the opinion x_i chosen by each agent i must belong to $[0,1] \cap [s_i - d_i^-, s_i + d_i^+]$, where s_i is the private belief of i, interval $[0,1]$ models the opinions space and d_i^-, d_i^+ are two given agent-specific thresholds.

As a consequence of her choice and of the choices of all the others, each agent i experiences a cost which depends on n functions: an increasing function g_i (*private cost function*), which measures the cost of agent i for disagreeing with her own belief, and $n-1$ functions $f_{i,j}$ for each $j \neq i$ (*public cost functions*), which measure the cost of the social pressure. In particular, $f_{i,j}$ is increasing (resp., decreasing) when agent j is a friend (resp., an enemy) of agent i. The cost experienced by agent i is formally defined as

$$c_i(x_1, \ldots, x_n) = \sum_{j \in [n]: j \neq i} f_{i,j}(|x_i - x_j|) + g_i(|x_i - s_i|),$$

where $x_j \in [0,1] \cap [s_j - d_j^-, s_j + d_j^+]$ is the opinion chosen by each agent j.

We stress that, despite of the huge mathematical challenges met in dealing with non-binary enemy relationships (one of the novelty of our model), most of our results only require all these functions to be continuous. Hence, our work provides a significant advancement along the direction of designing new models for opinion formation which may yield a good compromise between simplicity (needed for an analytical study) and expressive power.

Nevertheless, we also focus on special classes of games, that we name *well-ordered*, which turn out to enjoy interesting theoretical properties, while still spanning many realistic settings. Specifically, we consider opinion formation games that include the following additional properties: (i) the social groups do not intersect (and thus the opinions of the members of a group are always different from the opinions of the members of other groups), and (ii) all cost functions are strictly convex (i.e., the marginal increment of the cost strictly increases (resp., decreases) as the distance between opinions increases). The first property is realized when the social group membership is sufficiently strong to avoid any overlap of the opinions of agents belonging to different groups, despite they may influence each other. The second property is highly motivated in opinion dynamics, too. Indeed, convex cost functions model scenarios in which (a) the urgency of fixing the disagreement with close friends quickly grows as the disagreement becomes larger and larger, and similarly, (b) putting distance among enemies becomes more and more urgent when their opinions are close to each other. Furthermore, we point out that convexity is a common assumption in opinion formation games (see, e.g., [3, 10]), in which the influence functions are convex by hypothesis or coincide with some specific convex functions (e.g., quadratic or higher degree polynomials).

In light of the above considerations, our opinion formation framework and the special case of well-ordered games are able to include and generalize most of the previously defined models. Moreover, they can have multiple applications even in settings departing from opinion formation, such as facility location with heterogeneous preferences [16], content publishing [17] and isolation games [18, 19].

2. Our Contribution

We show that any game induced by our model admits at least a *pure Nash equilibrium* (i.e., a stable configuration in which each agent cannot reduce her cost via a unilateral change of opinion). We stress that this result does not require convexity or any other restrictive assumption to hold. In general, a game may admit different equilibria; however, we show that it is unique in well-ordered opinion formation games (that, differently from general games, must satisfy some convexity assumptions).

Next, we focus on the evaluation of the quality of equilibria through the concepts of *Price of Anarchy (PoA)* and *Price of Stability (PoS)*, by following the literature on the topic (see, e.g., [10, 11, 3, 5, 2, 20, 6, 8]. Indeed, PoA and PoS are used to better understand the social degradation caused by opinion formation phenomena that often appear in several real-life scenarios (e.g., political polls, trends formation, etc...). Moreover, PoA and PoS results play a practical role in establishing when the intervention of social planner is necessary, and when there is no need of altering the evolution of the system: whenever PoA/PoS are high, intervention of social planner may be welcome.

In this work, we focus on different ways to evaluating the quality of an equilibrium. A first approach uses the *utilitarian social cost*, defined as the sum of the agents' costs. This direction has been taken, e.g., in [3, 5, 10, 11]. A second approach emphasizes the *truthfulness* of the declared opinions, by bounding how much the social pressure deviates the agents' opinions from their private beliefs. This metric has been considered in [21, 22, 23]. A third approach, finally, measures the distance from a *consensus* [9, 13].

We believe that all these approaches are useful and meaningful. Not only, but it is often useful and meaningful to have, for example, equilibria that are close to be truthful (or close to be a consensus) and, at the same time, represent a good compromise for the society as a whole. For this reason, we propose to measure the performance of an equilibrium by means of the *λ-social influence cost*, obtained by summing the cost of untruthfulness scaled by λ and the cost of social pressure (i.e., distance from a consensus) scaled by $1-\lambda$, for any $\lambda \in [0,1]$. Observe that, by setting $\lambda = 1/2$, $\lambda = 1$ and $\lambda = 0$, respectively, we re-obtain the above three metrics.

Our results highlight how PoA and PoS with respect to λ-social influence cost vary as the parameters of the system change: this will provide practically useful suggestions about the direction in which possible interventions of a social planner should occur. For example, our results suggest that, in order to guarantee that opinion formation converges to states with good social performances, one should try to avoid enemy relation or one should try to assure that social groups are "closed" as described in the definition of well-ordered games. Hence, the social planner may be interested in designing campaigns to enforce these properties.

Specifically we prove that for extreme values of λ (i.e., $\lambda = 0$ or $\lambda = 1$), the PoA and the PoS can grow arbitrarily large, as it may be impossible to reach an equilibrium that is a consensus or a truthful profile when considering agents with general cost functions and possessing both attraction and dis-attraction attitudes. Nevertheless, we surprisingly show that the PoA and the PoS are usually not very large when λ is sufficiently far from the extremes. Specifically, we prove that the PoS is always (i.e., we do not require convexity or other assumptions) bounded by $\frac{\max\{2\lambda,1-\lambda\}}{\min\{2\lambda,1-\lambda\}} = O\left(\max\left\{\frac{1}{\lambda}, \frac{1}{1-\lambda}\right\}\right)$. The same bound holds even for the PoA in well-ordered

opinion formation games, while in general the PoA can be unbounded.

When the cost functions obey some additional mild assumptions (e.g., continuous differentiability), better bounds on the PoA are possible. In particular, given two classes of cost functions $\mathcal{F}$ and $\mathcal{G}$ satisfying the above assumptions, we show that the the price of anarchy $\mathsf{PoA}(\mathcal{O})$ (with respect to the λ-social influence) of any game $\mathcal{O}$ with public and private cost functions in $\mathcal{F}$ and $\mathcal{G}$ respectively is at most

$$\mathsf{PoA}(\mathcal{O}) \leq \inf_{\theta \geq 0} \left[\sup_{\substack{f \in \mathcal{F}, g \in \mathcal{G}, \\ x,y,\hat{x},\hat{y} \in [0,1]}} \max\left\{ \eta_{2\lambda}(\theta, f, x, y), \eta_{1-\lambda}(\theta, g, \hat{x}, \hat{y}) \right\} \right], \tag{1}$$

where $\eta_q(\theta, h, r, t)$ is a function defined as $\eta_q(\theta, h, r, t) = \frac{q \cdot h(r) + \theta(t-r) \cdot \frac{\partial h(r)}{\partial r}}{q \cdot h(t)}$ for any $\theta, q, r, t \geq 0$ and real function h.

The technique used to prove this result may be of independent interest: a generalization of the primal-dual technique introduced in [24], and applied for the first time in this setting. We additionally show that these bounds are often tight, even for the price of stability. Under mild assumptions, the proof arguments used to obtain the upper bound in (1) can be reversed via strong duality (by following a similar approach as in [25, 26, 27]) to derive tight lower bounds for the price of stability, holding even for games with two agents. The general structure of the lower bound is the following: (i) we have two agents with private beliefs equal to $1/2 - s$ and $1/2 + s$ for some $s \in [0, 1/2]$ (ii) there is a unique equilibrium $(x_1, x_2) = (1/2 - r, 1/2 + r)$ for some $r \in [0, 1/2]$ and the social optimum is $(y_1, y_2) = (1/2 - t, 1/2 + t)$ for some $t \in [0, 1/2]$.

We also apply the above results on the prices of anarchy and stability to specific classes of well-studied games, by proving tight numerical bounds.

References

[1] V. Bilò, D. Ferraioli, C. Vinci, General opinion formation games with social group membership, in: IJCAI, 2022, pp. 88–94.

[2] M. Degroot, Reaching a consensus, Journal of the American Statistical Association 69 (1974) 118–121.

[3] D. Bindel, J. M. Kleinberg, S. Oren, How bad is forming your own opinion?, Games and Economic Behavior 92 (2015) 248–265.

[4] N. Friedkin, E. Johnsen, Social influence and opinions, The Journal of Mathematical Sociology 15 (1990) 193–206.

[5] F. Chierichetti, J. Kleinberg, S. Oren, On discrete preferences and coordination, Journal of Computer and System Sciences 93 (2018) 11 – 29.

[6] D. Ferraioli, P. Goldberg, C. Ventre, Decentralized dynamics for finite opinion games, Theoretical Computer Science 648 (2016) 96–115.

[7] D. Fotakis, V. Kandiros, V. Kontonis, S. Skoulakis, Opinion dynamics with limited information, in: WINE, 2018, pp. 282–296.

[8] D. Fotakis, D. Palyvos-Giannas, S. Skoulakis, Opinion dynamics with local interactions, in: IJCAI, 2016, pp. 279–285.

[9] R. Hegselmann, U. Krause, Opinion dynamics and bounded confidence: Models, analysis and simulation, Journal of Artificial Societies and Social Simulation 5 (2002) 1–24.
[10] K. Bhawalkar, S. Gollapudi, K. Munagala, Coevolutionary opinion formation games, in: STOC, 2013, pp. 41–50.
[11] V. Bilò, A. Fanelli, L. Moscardelli, Opinion formation games with dynamic social influences, Theoretical Computer Science 746 (2018) 444–458.
[12] D. Ferraioli, C. Ventre, Social pressure in opinion games, in: IJCAI, 2017, pp. 3661–3667.
[13] V. Auletta, A. Fanelli, D. Ferraioli, Consensus in opinion formation processes in fully evolving environments, in: AAAI, 2019, pp. 6022–6029.
[14] V. Auletta, I. Caragiannis, D. Ferraioli, C. Galdi, G. Persiano, Generalized discrete preference games, in: IJCAI, 2016, pp. 53–59.
[15] E. Acar, G. Greco, M. Manna, Group reasoning in social environments, in: AAMAS, 2017, pp. 1296–1304.
[16] P. Serafino, C. Ventre, Heterogeneous facility location without money, Theoretical Computer Science 636 (2016) 27–46.
[17] V. Bilò, M. Flammini, C. Vinci, The quality of content publishing in the digital era, in: ECAI, 2020, pp. 35–42.
[18] V. Bilò, M. Flammini, G. Monaco, L. Moscardelli, On the performances of nash equilibria in isolation games, Journal of Combinatorial Optimization 22 (2011) 378 – 391.
[19] Y. Zhao, W. Chen, S. Teng, The isolation game: A game of distances, in: ISAAC, 2008, pp. 148–159.
[20] M. Epitropou, D. Fotakis, M. Hoefer, S. Skoulakis, Opinion formation games with aggregation and negative influence, Theory Comput. Syst. 63 (2019) 1531–1553.
[21] V. Auletta, I. Caragiannis, D. Ferraioli, C. Galdi, G. Persiano, Minority becomes majority in social networks, in: WINE, 2015, pp. 74–88.
[22] V. Auletta, I. Caragiannis, D. Ferraioli, C. Galdi, G. Persiano, Information retention in heterogeneous majority dynamics, in: WINE, 2017, pp. 30–43.
[23] V. Auletta, D. Ferraioli, G. Greco, Reasoning about consensus when opinions diffuse through majority dynamics, in: IJCAI, 2018, pp. 49–55.
[24] V. Bilò, A unifying tool for bounding the quality of non-cooperative solutions in weighted congestion games, Theory of Computing Systems 62 (2018) 1288 – 1317.
[25] V. Bilò, C. Vinci, On the impact of singleton strategies in congestion games, in: ESA, 2017, pp. 17:1–17:14.
[26] V. Bilò, L. Moscardelli, C. Vinci, Uniform mixed equilibria in network congestion games with link failures, in: ICALP, 2018, pp. 146:1–146:14.
[27] F. Benita, V. Bilò, B. Monnot, G. Piliouras, C. Vinci, Data-driven models of selfish routing: Why price of anarchy does depend on network topology, in: WINE, 2020, pp. 252–265.