Minority Becomes Majority in Social Networks

Vincenzo Auletta[1], Ioannis Caragiannis[2], Diodato Ferraioli[1](✉), Clemente Galdi[3], and Giuseppe Persiano[1]

[1] Università degli Studi di Salerno, Fisciano, Italy
{auletta,dferraioli,pino.persiano}@unisa.it
[2] CTI "Diophantus" and University of Patras, Patras, Greece
caragian@ceid.upatras.gr
[3] Università di Napoli "Federico II", Napoli, Italy
clemente.galdi@unina.it

Abstract. It is often observed that agents tend to imitate the behavior of their neighbors in a social network. This imitating behavior might lead to the strategic decision of adopting a public behavior that differs from what the agent believes is the right one and this can subvert the behavior of the population as a whole.

In this paper, we consider the case in which agents express preferences over two alternatives and model social pressure with the *majority* dynamics: at each step an agent is selected and its preference is replaced by the majority of the preferences of her neighbors. In case of a tie, the agent does not change her current preference. A profile of the agents' preferences is *stable* if the each agent's preference coincides with the preference of at least half of the neighbors (thus, the system is in equilibrium).

We ask whether there are network topologies that are robust to social pressure. That is, we ask whether there are graphs in which the majority of preferences in an initial profile $\mathbf{s}$ always coincides with the majority of the preference in all stable profiles reachable from $\mathbf{s}$. We completely characterize the graphs with this robustness property by showing that this is possible only if the graph has no edge or is a clique or very close to a clique. In other words, except for this handful of graphs, every graph admits at least one initial profile of preferences in which the majority dynamics can subvert the initial majority. We also show that deciding whether a graph admits a minority that becomes majority is NP-hard when the minority size is at most 1/4-th of the social network size.

This work was partially supported by the COST Action IC1205 "Computational Social Choice", by the Italian MIUR under the PRIN 2010–2011 project ARS TechnoMedia – Algorithmics for Social Technological Networks, by the European Social Fund and Greek national funds through the research funding program Thales on "Algorithmic Game Theory", by the EU FET project MULTIPLEX 317532, and by a Caratheodory basic research grant from the University of Patras.

E. Markakis and G. Schäfer (Eds.): WINE 2015, LNCS 9470, pp. 74–88, 2015.
DOI: 10.1007/978-3-662-48995-6_6

1 Introduction

Social scientists are greatly interested in understanding how social pressure can influence the behavior of agents in a social network. We consider the case in which agents connected through a social network must choose between two alternatives and, for concreteness, we consider two competing technologies: the current (or old) technology and a new technology. To make their decision, the agents take into account two factors: their personal relative valuation of the two technologies and the opinions expressed by their social neighbors. Thus, the public action taken by an agent (i.e., adopting the new technology or staying with the old) is the result of a mediation between her personal valuation and the social pressure derived from her neighbors.

The first studies concerning the adoption of new technologies date back to the middle of 20-th century, with the analysis of the adoption of hybrid seed corn among farmers in Iowa [16] and of tetracycline by physicians in US [6].

We assume that agents receive an initial signal about the quality of the new technology that constitutes the agent's initial preference. This signal is independent from the agent's social network; e.g., farmers acquired information about the hybrid corn from salesman and physicians acquired information about tetracycline from scientific publications. After the initial preference is formed, an agent tends to conform her preference to the one of her neighbors and thus to *imitate* their behavior, even if this disagrees with her own initial preference. This imitating behavior can be explained in several ways: an agent that sees a majority agreeing on an opinion might think that her neighbors have access to some information unknown to her and hence they have made the better choice; also agents can directly benefit from adopting the same behavior as their friends (e.g., prices going down).

Thus, the natural way of modeling the evolution of preferences in networks is through a majority dynamics: each agent has an initial preference and at each time step a subset of agents updates their opinion conforming to the majority of their neighbors in the network. As a tie-breaking rule it is usual to assume that when exactly half of the neighbors adopted the new technology, the agent decides to stay with her current choice to avoid the cost of a change. Thus, the network undergoes an opinion formation process where agents continue to update their opinions until a stable profile is reached, where each agent's behavior agrees with the majority of her neighbors. Notice that the dynamics does not take into account the relative merits of the two technologies and, without loss of generality, we adopt the convention that the technology that is preferred by the majority of the agents in the initial preference profile is the new technology.

In the setting described above, it is natural to ask whether and when the social pressure of conformism can change the opinion of some of the agents so that the initial majority is subverted. In the case of the adoption of a new technology, we are asking whether a minority of agents supporting the old technology can orchestrate a campaign and convince enough agents to reject the new technology, even if the majority of the agents had initially preferred the new technology.

This problem has been extensively studied in the literature. If we assume that updates occur *sequentially*, one agent at each time step, then it is easy to design graphs (e.g., a star) where the old technology, supported by an arbitrarily small minority of agents, can be adopted by most of the agents. Berger [2] proved that such a result holds even if at each time step all agents *concurrently* update their actions. However, Mossel et al. [13] and Tamuz and Tessler [17] proved that there are graphs for which, both with concurrent and sequential updates, at the end of the update process the new technology will be adopted by the majority of agents with high probability.

In [9,13] it is also proved that when the graph is an expander, agents will reach a *consensus* on the new technology with high probability for both sequential and concurrent updates (the probability is taken on the choice of initial configurations with a majority of new technology adopters). Thus, expander graphs are particularly efficient in aggregating opinions since, with high probability, social pressure does not prevent the diffusion of the new technology.

In this paper, we will extend this line of research by taking a worst-case approach instead of a probabilistic one. We ask whether there are graphs that are robust to social pressure, even when it is driven by a carefully and adversarially designed campaign. Specifically, we want to find out whether there are graphs in which no subset of the agents preferring the old technology (and thus consisting of less than half of the agents) can manipulate the rest of the agents and drive the network to a stable profile in which the majority of the agents prefers the old technology. This is easily seen to hold for two extreme graphs: the clique and the graph with no edge. In this paper, we prove that these are essentially[1] the only graphs where social pressure cannot subvert the majority.

In particular, our results highlight that even for expander graphs, where it is known that agents converge with high probability to consensus on the new technology, it is possible to fix a minority and orchestrate a campaign that brings the network into a stable profile where at least half of the agents decide to not adopt the new technology.

Overview of our contribution. We consider the following sequential dynamics. We have n agents and at any given point the system is described by the profile $\mathbf{s}$ in which $\mathbf{s}(i) \in \{0, 1\}$ is the preference of the i-th agent. We say that agent i is *unhappy* in profile $\mathbf{s}$ if the majority of her neighbors have a preference different from $\mathbf{s}(i)$. Profiles evolve according to the dynamics in which an *update* consists of non-deterministically selecting an unhappy agent and changing its preference. A profile in which no agent is unhappy is called *stable*.

In Sect. 2 (see Theorems 1 and 2), we characterize the set of social networks (graphs) where a majority can be subverted by social pressure. More specifically, we show that for each of these graphs it is possible to select a minority of agents not supporting the new technology and a sequence of updates (a campaign) that leads the network to a stable profile where the majority of the agents prefers the old technology. As described above, we will prove that this class is very large

[1] It turns out that for an even number of nodes, there are a few more very dense graphs enjoying such a property.

and contains all graphs except a small set of forbidden graphs, consisting of the graph with no edges and of other graphs that are almost cliques. Proving this fact turned out to be a technically challenging task and it is heavily based on properties of local optima of graph bisections.

Then we turn our attention to related computational questions. First we show that we can compute in polynomial time an initial preference profile, where the majority of the agents supports the new technology, and a sequence of update that ends in a stable profile where at least half of the agents do not adopt the new technology. This is done through a polynomial-time local-search computation of a bisection of locally minimal width.

We actually prove a stronger result. In principle, it could be that from the starting profile the system needs to undergo a long sequence of updates, in which the minority gains and loses member to eventually reach a stable profile in which the minority has become a majority. Our algorithm shows that this can always be achieved by means of a short sequence of at most two updates after which any sequence of updates will bring the system to a stable profile in which the initial minority has become majority. This makes the design of an adversarial campaign even more realistic, since such a campaign only has to identify the few "swing" agents and thus it turns out to be very simple to implement.

However, the simplicity of the subverting campaign comes at a cost. Indeed, our algorithm always computes an initial preferences profile that has very large minorities, consisting of $\left\lfloor \frac{n-1}{2} \right\rfloor$ agents. We remark that, even in case of large minorities, it is not trivial to give a sequence of update steps that ends in a stable profile where the majority is subverted. Indeed, even if the large minority of the original profile makes it easy to find a few agents of the original majority that prefer to change their opinions, this is not sufficient in order to prove that the majority has been subverted, since we have also to prove that there are no other nodes in the original minority that prefer to change their preference.

Moreover, we observe that, even if there are cases in which such a large minority is necessary, the idea behind our algorithm can be easily turned into an heuristic that checks whether the majority can be subverted by a smaller minority (e.g., by considering unbalanced partitions in place of bisections).

On the other side, we show that a large size of the minority in the initial preference profile seems to be necessary in order to quickly compute a subverting minority and its corresponding sequence of updates. Indeed, given a n-node social network, deciding whether there exists a minority of less than $n/4$ nodes and a sequence of update steps that bring the system to a stable profile in which the majority has been subverted is an NP-hard problem (see Theorem 4).

The main source of computational hardness seems to arise from the computation of the initial preference profile. Indeed, if this profile is given, computing the maximum number of adopters of the new technology (and, hence, deciding whether majority can be subverted) and the corresponding sequence of updates turns out to be possible in polynomial time (see Theorem 5).

Related work. There is a vast literature on the effect that social pressure has on the behavior of a system as a whole. In many works, influence is modeled by agents simply following the majority [2,9,13,17]. A generalization of this imitating behavior is discussed in [13].

A different approach is taken in [14], where each agent updates her behavior according to a Bayes rule that takes in account its own initial preference and what is declared by neighbors on the network.

Yet another approach assumes that agents are strategic and rational. That is, they try to maximize some utility function that depends on the level of coordination with the neighbors on the network. Here, the updates occur according to a best response dynamics or some other more complex game dynamics. Along this direction, particularly relevant to our works are the ones considering best-response dynamics from truthful profiles in the context of iterative voting, e.g., see [4,12]. In particular, closer to our current work is the paper of Brânzei et al. [4] who present bounds on the quality of equilibria that can be reached from a truthful profile using best-response play and different voting rules. The important difference is that there is no underlying network in their work.

Our work is also strictly related with a line of work in social sciences that aims to understand how opinions are formed and expressed in a social context. A classical simple model in this context has been proposed by Friedkin and Johnsen [11] (see also [7]). Its main assumption is that each individual has a private initial belief and that the opinion she eventually expresses is the result of a repeated averaging between her initial belief and the opinions expressed by other individuals with whom she has social relations. The recent work of Bindel et al. [3] assumes that initial beliefs and opinions belong to $[0, 1]$ and interprets the repeated averaging process as a best-response play in a naturally defined game that leads to a unique equilibrium.

An obvious refinement of this model is to consider discrete initial beliefs and opinions by restricting them, for example, to two discrete values (see [5,10]). Clearly, the discrete nature of the opinions does not allow for averaging anymore and several nice properties of the opinion formation models mentioned above — such as the uniqueness of the outcome — are lost. In contrast, in [10] and in [5], it is assumed that each agent is strategic and aims to pick the most beneficial strategy for her, given her internal initial belief and the strategies of her neighbors. Interestingly, it turns out that the majority rule used in this work for describing how agents update their behavior can be seen as a special case of the discrete model of [5,10], in which agents assign a weight to the initial preference smaller than the one given to the opinion of the neighbors.

Studies on social networks consider several phenomena related to the spread of social influence such as information cascading, network effects, epidemics, and more. The book of Easley and Kleinberg [8] provides an excellent introduction to the theoretical treatment of such phenomena. From a different perspective, problems of this type have also been considered in the distributed computing literature, motivated by the need to control and restrict the influence of failures in distributed systems; e.g., see the survey by Peleg [15] and the references therein.

Preliminaries. We formally describe our model as follows. There are n agents; we use $[n] = \{1, 2, ..., n\}$ to denote their set. Each agent corresponds to a distinct node of a graph $G = (V, E)$ that represents the *social network*; i.e., the network of social relations between the agents. Agent i has an initial preference $\mathbf{s}_0(i) \in \{0, 1\}$. At each time step, agent i can update her preference to $\mathbf{s}(i) \in \{0, 1\}$. A *profile* is a vector of preferences, with one preference per agent. We use bold symbols for profiles; i.e., $\mathbf{s} = (\mathbf{s}(1), \ldots, \mathbf{s}(n))$. In particular, we sometimes call the profile of initial preferences $(\mathbf{s}_0(1), \ldots, \mathbf{s}_0(n))$ as the *truthful profile*. Moreover, for any $y \in \{0, 1\}$, we denote as $\overline{y}$ the negation of y; i.e., $\overline{y} = 1 - y$.

A graph G is mbM (*minority becomes majority*) if there exists a profile $\mathbf{s}_0$ of initial preferences such that: the number of nodes that prefer 0 is a strict majority, i.e., $|\{x \in V : \mathbf{s}_0(x) = 0\}| > n/2$; and there is a *subverting* sequence of updates that starts from $\mathbf{s}_0$ and reaches a stable profile $\mathbf{s}$ in which the number of nodes that prefer 0 is not a majority, i.e., $|\{x \in V : \mathbf{s}(x) = 0\}| \leq n/2$. A profile of initial preferences that witnesses a graph being mbM will be also termed mbM.

2 Characterizing the mbM Graphs

The main result of this section is a characterization of the mbM graphs. More formally, we have the following definition. A graph G with n nodes is *forbidden* if one of the following conditions is satisfied:
F1: G has no edge;
oF2: G has an odd number of nodes, all of degree $n - 1$ (that is, G is a clique);
eF2: G has an even number of nodes and all its nodes have degree at least $n - 2$;
eF3: G has an even number of nodes, $n - 1$ nodes of G form a clique, and the remaining node has degree at most 2;
eF4: G has an even number of nodes, $n - 1$ nodes of G have degree $n - 2$ but they do not form a clique, and the remaining node has degree at most 4.

We begin by proving the following statement.

Theorem 1. *No forbidden graph is* mbM.

Proof. We will distinguish between cases for a forbidden graph G. Clearly, if G is F1, then it is not mbM since no node can change its preference. Now assume that G is eF2 (respectively, oF2) and consider a profile in which there are at least $\frac{n}{2} + 1$ (respectively, $\frac{n+1}{2}$) agents with preference 0. Then, every node x with initial preference 0 has at most $\frac{n}{2} - 1$ neighbors with initial preference 1 and at least $\frac{n}{2} - 1$ neighbors with initial preference 0 (respectively, at most $\frac{n-1}{2}$ neighbors with initial preference 1 and at least $\frac{n-1}{2}$ neighbors with initial preference 0). Hence, x is not unhappy and stays with preference 0.

Now, consider the case where G is eF3 and let u be the node of degree at most 2. Consider profile $\mathbf{s}_0$ of initial preferences in which there are at least $\frac{n}{2} + 1$ agents with preference 0. First observe that in the truthful profile $\mathbf{s}_0$ any node x other than u that has preference 0 is adjacent to at most $\frac{n}{2} - 1$ nodes with initial preference 1 and to at least $\frac{n}{2} - 1$ nodes with initial preference 0. Then, x is not unhappy and stays with preference 0. Hence, u is the only node that may want to switch from 0 to 1. But this is possible only if all nodes in the neighborhood of u have preference 1, which implies that the neighborhood of any node with initial preference 0 does not change after the switch of u, i.e., nodes with preference 0 still are not unhappy and thus they have no incentive to switch to 1. Then, any node with preference 1 that is not adjacent to u has at most $\frac{n}{2} - 2$ neighbors with preference 1 and at least $\frac{n}{2}$ neighbors with preference 0. Also, any node with preference 1 that is adjacent to u has $\frac{n}{2} - 1$ neighbors with preference 1 and $\frac{n}{2}$ neighbors with preference 0. So, every node with preference 1 will eventually switches to 0.

It remains to consider the case where G is eF4; let u be the node of degree at most 4. Actually, it can be verified that u can have degree either 2 or 4 and its neighbors form pair(s) of non-adjacent nodes. Consider a truthful profile in which there are at least $\frac{n}{2} + 1$ agents with preference 0. Observe that a node different from u that has initial preference 0 has at most $\frac{n}{2} - 1$ neighbors with preference 1 and at least $\frac{n}{2} - 1$ neighbors with preference 0. So, it is not unhappy and has no incentive to switch to preference 1. The only node that might do so is u, provided that the strict majority of its neighbors (i.e., both of them if u has degree 2 and at least three of them if u has degree 4) have preferences 1. This switch cannot trigger another switch of the preference of an agent from 0 node to 1. Indeed, there is at most one agent with preference 0 that can be adjacent to u. Since this node is not adjacent to one of the neighbors of u with preference 1, it has at most $\frac{n}{2} - 1$ neighbors with preference 1 (and at least $\frac{n}{2} - 1$ neighbors with preference 0). Hence, it has no incentive to switch to preference 1 either. Now, consider two neighbors of u with preference 1 that are not adjacent (these nodes certainly exist). Each of them is adjacent to $\frac{n}{2} - 2$ nodes with preference 1 and $\frac{n}{2}$ nodes with preference 0. Hence, they have an incentive to switch to 0. Then, the number of nodes with preference 1 is at most $\frac{n}{2} - 2$ and eventually all nodes will switch to preference 0. □

The following is the main result of this section.

Theorem 2. *Every non-forbidden graph is* mbM.

We next give the proof for the simpler case of graphs with an odd number of nodes and postpone the full proof to the full version [1]. Let us start with the following definitions. A *bisection* $\mathcal{S} = (S, \overline{S})$ of a graph $G = (V, E)$ with n nodes is simply a partition of the nodes of V into two sets S and $\overline{S}$ of sizes $\lceil n/2 \rceil$ and $\lfloor n/2 \rfloor$, respectively. We will refer to S and $\overline{S}$ as the *sides* of bisection $\mathcal{S}$. The *width* $W(S, \overline{S})$ of a bisection $\mathcal{S}$ is the number of edges of G whose endpoints belong to different sides of the partition. The *minimum* bisection $\mathcal{S}$ of G has minimum width among all partitions of G. We extend notation $W(A, B)$ to any pair (A, B) of subsets of nodes of G in the obvious way. When $A = \{x\}$ is a singleton we will write $W(x, B)$ and similarly for B. Thus, if nodes x and y are adjacent, then $W(x, y) = 1$; otherwise $W(x, y) = 0$. For a bisection $\mathcal{S} = (S, \overline{S})$, we define the *deficiency* $\mathsf{def}_{\mathcal{S}}(x)$ *of node x w.r.t. bisection* $\mathcal{S}$ as $\mathsf{def}_{\mathcal{S}}(x) = W(x, S) - W(x, \overline{S})$ if $x \in S$, and $\mathsf{def}_{\mathcal{S}}(x) = W(x, \overline{S}) - W(x, S)$ if $x \in \overline{S}$. Let $\mathcal{S} = (S, \overline{S})$ be a minimum bisection of a graph G with n nodes. Then it is not hard to see that minimality of $\mathcal{S}$ implies that for every $x \in S$ and $y \in \overline{S}$, $\mathsf{def}_{\mathcal{S}}(x) + \mathsf{def}_{\mathcal{S}}(y) + 2W(x, y) \geq 0$. Moreover if n is odd, $\mathsf{def}_{\mathcal{S}}(x) \geq 0$.

We have the following technical lemma.

Lemma 1. *Suppose that a graph G admits a bisection $\mathcal{S} = (S, \overline{S})$ in which S consists of nodes with non-negative deficiency and includes at least one node with positive deficiency. Then G is* mbM.

Proof. Let v be the node with positive deficiency in S and consider profile $\mathbf{s}_0$ of initial preferences in which any node in S except v has preference 1 and remaining nodes have preference 0. Hence, in $\mathbf{s}_0$ there is a majority of $\lceil n/2 \rceil$ agents with preference 0. Observe also that in $\mathbf{s}_0$, v is adjacent to $W(v, S)$ nodes with preference 1 and to $W(v, \overline{S})$ nodes with preference 0. Since $\mathsf{def}_{\mathcal{S}}(v) > 0$ then v is unhappy with preference 0 and updates her preference to 1. We thus reach a profile $\mathbf{s}_1$ in which $\lceil n/2 \rceil$ nodes have preference 1 (that is, all nodes in S). We conclude the proof of the lemma by showing that every node of S is not unhappy and thus it stays with preference 1[2]. This is obvious for v. Let us consider $u \in S$ and $u \neq v$. Then u has $W(u, S)$ neighbors with preference 1 and $W(u, \overline{S})$ neighbors with preference 0. Since $\mathsf{def}_{\mathcal{S}}(u) \geq 0$, we have that $W(u, S) \geq W(u, \overline{S})$. Hence, the number of neighbors of u with preference 0 is not a majority. Then, u is not unhappy, and thus stays with preference 1. □

We are now ready to prove Theorem 2 for odd-sized graphs. We remind the reader that the (more complex) proof for even-size graphs is in the full version [1].

Proposition 3. *Non-forbidden graphs with an odd number of nodes are* mbM.

Proof. Let G be a non-forbidden graph with an odd number of nodes and let $\mathcal{S} = (S, \overline{S})$ be a minimum bisection for G. By minimality of $\mathcal{S}$, we have that $\mathsf{def}_{\mathcal{S}}(x) \geq 0$, for all $x \in S$. If S contains at least a node v with $\mathsf{def}_{\mathcal{S}}(v) > 0$ then, by minimality of $\mathcal{S}$, G is mbM. So assume that $\mathsf{def}_{\mathcal{S}}(x) = 0$ for all $x \in S$.

Minimality of $\mathcal{S}$ implies that if $\mathsf{def}_{\mathcal{S}}(v) < 0$ for $v \in \overline{S}$ then $\mathsf{def}_{\mathcal{S}}(v) \geq -2$ and v is connected to all nodes in S. Therefore $W(v, S) = \lceil n/2 \rceil$ and, since $W(v, \overline{S}) \leq \lfloor n/2 \rfloor - 1$, $\mathsf{def}_{\mathcal{S}}(v) = -2$. We denote by A the set of all the nodes $y \in \overline{S}$ with $\mathsf{def}_{\mathcal{S}}(x) = -2$; therefore, all nodes $y \in \overline{S} \setminus A$ have $\mathsf{def}_{\mathcal{S}}(y) \geq 0$.

Let us first consider the case in which $A \neq \emptyset$ and there are two non-adjacent nodes $u, w \in S$. Then pick any node $v \in A$ and consider partition $\mathcal{T} = (T, \overline{T})$ with $T = S \cup \{v\} \setminus \{u\}$. We have that $W(v, T) = W(v, S) - 1 = \lceil n/2 \rceil - 1$ and $W(v, \overline{T}) = W(v, \overline{S}) + 1 = \lfloor n/2 \rfloor + 1$ and hence $\mathsf{def}_{\mathcal{T}}(v) = 0$. For any $x \in T \setminus \{v, w\}$, we have $\mathsf{def}_{\mathcal{T}}(x) \geq \mathsf{def}_{\mathcal{S}}(x) = 0$. Node w is connected to v but not to u and, thus, $\mathsf{def}_{\mathcal{T}}(w) \geq \mathsf{def}_{\mathcal{S}}(w) + 2 = 2$. Then, by Lemma 1, G is mbM.

Assume now that $A \neq \emptyset$ and S is a clique. That is, $W(x, S) = \lceil n/2 \rceil - 1$ for every $x \in S$, and, since $\mathsf{def}_{\mathcal{S}}(x) = 0$, it must be that $W(x, \overline{S}) = W(x, S)$ and thus x is connected to all nodes in $\overline{S}$. Therefore, for all $y \in \overline{S}$, $W(y, S) = \lceil n/2 \rceil$

[2] This is sufficient since the switch of nodes in $\overline{S}$ that are unhappy with preference 0 only increases the number of nodes with preference 1. Moreover, if some nodes in $\overline{S}$ switch their preferences, then the number of nodes with preference 1 in the neighborhood of any node in S can only increase.

and, since $\mathsf{def}_{\mathcal{S}}(y) \geq -2$ it must be that $W(y, \overline{S}) \geq \lceil n/2 \rceil - 2 = |\overline{S}| - 1$. In other words, every node of $\overline{S}$ is connected to every node of $\overline{S}$ and thus G is a clique.

Finally, assume that $A = \emptyset$; that is, $\mathsf{def}_{\mathcal{S}}(y) \geq 0$ for any $y \in \overline{S}$. If for some $v \in \overline{S}$, we have $\mathsf{def}_{\mathcal{S}}(v) > 0$, then consider partition $\mathcal{T} = (T, \overline{T})$ with $T = \overline{S} \cup \{u\}$, where u is any node from S. For any $x \in T \cap \overline{S}$, $\mathsf{def}_{\mathcal{T}}(x) \geq \mathsf{def}_{\mathcal{S}}(x) \geq 0$, $\mathsf{def}_{\mathcal{T}}(u) = -\mathsf{def}_{\mathcal{S}}(u) = 0$ and $\mathsf{def}_{\mathcal{T}}(v) \geq \mathsf{def}_{\mathcal{S}}(v) \geq 1$. By Lemma 1, G is mbM.

Finally, we consider the case in which $\mathsf{def}_{\mathcal{S}}(y) = 0$ for every node x of G. Since G is not empty, there exists at least one edge in G and, since the endpoints of this edge have $\mathsf{def}_{\mathcal{S}} = 0$ there must be at least node $v \in S$ with a neighbor $w \in \overline{S}$. Now, consider partition $\mathcal{T} = (T, \overline{T})$ with $T = S \cup \{w\}$. We have that every node $x \in T \cap S$, has $\mathsf{def}_{\mathcal{T}}(x) \geq \mathsf{def}_{\mathcal{S}}(x) = 0$, $\mathsf{def}_{\mathcal{T}}(w) = -\mathsf{def}_{\mathcal{S}}(w) = 0$, and $\mathsf{def}_{\mathcal{T}}(w) > \mathsf{def}_{\mathcal{S}}(w) = 0$. The claim again follows by Lemma 1. □

We note that the only property required is local minimality. Since a local-search algorithm can compute a locally minimal bisection in polynomial time, we can make constructive the proof of Proposition 3, and quickly compute the subverting minority and the corresponding updates.

3 Hardness for Weaker Minorities

We next show that deciding if it is possible to subvert the majority starting from a weaker minority is a computationally hard problem.

Theorem 4. *For every constant $0 < \varepsilon < \frac{1}{8}$, given a graph G with n nodes, it is NP-hard to decide whether there exists an* mbM *profile of initial preferences with at most $n\left(\frac{1}{4} - \varepsilon\right)$ nodes with initial preference* 1*.*

Proof. We will use a reduction from the NP-hard problem 2P2N-3SAT, the problem of deciding whether a 3SAT formula in which every variable appears as positive in two clauses and as negative in two clauses has a truthful assignment or not (the NP-hardness follows by the results of [18]).

Given a Boolean formula ϕ with C clauses and V variables that is an instance of 2P2N-3SAT (thus $3C = 4V$ and C is a multiple of 4), we will construct a graph $G(\phi)$ with n nodes such that there exists a profile of initial preferences with at most $n\left(\frac{1}{4} - \varepsilon\right)$ nodes of $G(\phi)$ with preference 1 such that a sequence of updates can lead to a stable profile in which at least $n/2$ nodes have preference 1 if and only if ϕ has a satisfying assignment.

The graph $G(\phi)$ has the following nodes and edges. For each variable x of ϕ, $G(\phi)$ includes a *variable* gadget for x consisting of 25 nodes and 50 edges. The nodes of the variable gadget for x are the *literal nodes*, x and $\overline{x}$, nodes $v_1(x), \ldots, v_7(x)$, nodes $v_1(\overline{x}), \ldots, v_7(\overline{x})$, nodes $v_0(x)$ and $w_0(x)$, and nodes $w_1(x)$, $\ldots, w_7(x)$. The edges are $(x, v_i(x))$ and $(\overline{x}, v_i(\overline{x}))$ for $i = 1, \ldots, 7$, $(v_i(x), v_{i+1}(x))$ and $(v_i(\overline{x}), v_{i+1}(\overline{x}))$ for $i = 1, \ldots, 6$, $(v_0(x), v_7(x))$, $(v_0(x), v_7(\overline{x}))$, $(v_0(x), w_0(x))$, $(w_0(x), v_i(x))$, $(w_0(x), v_i(\overline{x}))$ and $(w_0(x), w_i(x))$ for $i = 1, \ldots, 7$. For each clause c of ϕ, graph $G(\phi)$ includes a *clause gadget* for c consisting of 18 nodes and 32 edges. The nodes of the gadget are the *clause* node c, nodes $u_1(c)$, $u_2(c)$, and nodes $v_1(c), \ldots, v_{15}(c)$. The 32 edges are $(c, u_1(x)), (c, u_2(x))$, and $(u_i(c), v_j(c))$ with $i = 1, 2$ and $j = 1, \ldots, 15$. In $G(\phi)$, for every clause c, the clause node c is connected to the three literal nodes corresponding to the literals that appear in clause c in ϕ. Therefore, each literal node is connected to the two clauses in which it appears. Graph $G(\phi)$ includes a *clique* of even size N, with $12C \leq N \leq \frac{95C}{16\varepsilon} - \frac{123C}{4}$; the clique is disconnected from the rest of the graph. Graph $G(\phi)$ includes $N + \frac{99C}{4}$ additional *isolated* nodes. Overall, the total number of nodes in $G(\phi)$ is $n = 2N + \frac{99C}{4} + 25V + 18C = 2N + \frac{123C}{2}$.

A profile of initial preferences to the nodes of $G(\phi)$ is called *proper* if: for every variable x, it assigns preference 1 to node $w_0(x)$ and to exactly one literal node of the gadget of x; for every clause c, it assigns preference 1 to nodes $u_1(c)$ and $u_2(c)$; it assigns preference 1 to exactly $\frac{N}{2}$ nodes of the clique; it assigns preference 0 to all the remaining nodes. Hence, in a proper profile the number of nodes with preference 1 is $2V + 2C + \frac{N}{2} = \frac{7C}{2} + \frac{N}{2} \leq n(\frac{1}{4} - \varepsilon)$; the inequality follows from the upper bound in the definition of N.

We now prove that $G(\phi)$ has a proper profile of initial preferences that leads to a majority of nodes with preference 1 if and only ϕ is satisfiable. First observe that every clique node switches her preference to 1 (as the strict majority of its neighbors has initially preference 1 and this number gradually increases until all clique nodes switch to 1).

We next prove that starting from a proper profile of initial preferences, there is a sequence of updates that leads to a stable profile in which 17 nodes of every variable gadget have preference 1. To see this, consider a proper profile that assigns preference 1 to x (and to $w_0(x)$) and the following sequence of updates: node $v_1(x)$ switches from 0 to 1; then, for $i = 1, \ldots, 6$, node $v_{i+1}(x)$ switches to 1 immediately after node $v_i(x)$; node $v_0(x)$ switches to 1 after node $v_7(x)$; finally, $w_1(x), \ldots, w_7(x)$ can switch in any order. Observe that in this sequence any switching node is unhappy since it has a strict majority of nodes with preference 1 in its neighborhood. Also, the resulting profile where the 17 nodes $w_0(x), w_1(x), \ldots, w_7(x)$, $v_0(x), v_1(x), \ldots, v_7(x)$, and x have preference 1 is stable, i.e., no node in the gadget is unhappy. Indeed, for each node with preference 1, the strict majority of the preferences of its neighbors is 1. Hence, the node has no incentive to switch to preference 0. For each of the remaining nodes (with preference 0), at least half of its neighbors is 0. Hence, this node has no incentive to switch to preference 1 either. A similar sequence can be constructed for a proper profile that assigns preference 1 to node $\overline{x}$ (and $w_0(x)$) of the gadget for variable x. Intuitively, the two proper profiles of initial preferences simulate the assignment of values TRUE and FALSE to variable x, respectively.

In addition, it is easy to see that starting from a proper profile, there is no sequence of updates that reaches a stable profile where more than 17 nodes in a variable gadget have preference 1 (the observation needed here is the same that guarantees that we reach a stable profile above).

Let us now consider the clause gadgets associated with clause c of ϕ. We observe that, starting from a proper profile of initial preferences, there exists a sequence of updates that leads to a stable profile in which 17 nodes of the clause gadget have preference 1. Indeed, starting from the proper assignment of preference 1 to nodes $u_1(c)$ and $u_2(c)$, nodes $v_1(c), \ldots, v_{15}(c)$ will switch from 0 to 1 in arbitrary order (for each of them both neighbors have preference 1). After these updates, at least 15 out of 17 neighbors of $u_1(c)$ and $u_2(c)$ have preference 1 and both neighbors of the nodes $v_1(c), \ldots, v_{15}(c)$ have preference 0. Hence, none among these nodes have any incentive to switch to preference 0.

Let us now focus on the clause nodes and observe that node c in the corresponding clause gadget will switch to 1 if and only if at least one of the literal nodes corresponding to literals that appear in c have preference 1 (since the degree of a clause node is five and nodes $u_1(c)$ and $u_2(c)$ have preference 1). This switch cannot trigger any other switch in literal nodes or in nodes of clause gadgets since the preference of these nodes coincides with a strict majority of preferences in its neighborhood. Hence, the fact that a clause node has preference 1 (respectively, 0) corresponds to the clause being satisfied (respectively, not satisfied) by the Boolean assignment induced by the proper profile of initial preferences. Eventually, the updates lead to an additional number of C clause nodes adopting preference 1 in the stable profile if and only if ϕ is satisfiable.

In conclusion, we have that if ϕ is satisfiable there is a sequence of updates converging to a profile with $17V + 17C + N + C = N + 123C/4 = n/2$ nodes with preference 1. Otherwise, if ϕ is not satisfiable, any sequence of updates converges to a stable profile with strictly less than $n/2$ nodes having preference 1.

We conclude the proof by showing that it is sufficient to restrict to proper assignments as non-proper assignments will never lead to a stable profile with a majority of nodes with preference 1. First observe that if the total number of clique and isolated nodes with preference 1 is strictly less than $\frac{N}{2}$, then no clique and isolated nodes with preference 0 will adopt preference 1. Thus, in this case, even counting all nodes in variable and clause gadgets, any sequence of updates converges to a stable profile with at most $25V + 18C + \frac{N}{2} - 1 < \frac{n}{2}$ nodes with preference 1 (where we used that $N \geq 12C$).

Let us now focus on a profile of initial preferences that assigns preference 1 to at most $7C/2 = 2C + 2V$ nodes from variable and clause gadgets. Suppose that this profile of initial preferences is such that a sequence of updates leads to a stable profile with at least $n/2$ nodes with preference 1. We will show that this profile of initial preferences must be proper.

First, observe that if at most one node in a variable gadget is assigned preference 1, then all nodes in the gadget will eventually adopt preference 0 after a sequence of updates. Indeed, a literal node will have at least six neighbors with preference 0 and at most three with preference 1, and any non-literal node will have at most one out of at least three of its neighbors with preference 0.

Consider now profiles of initial preferences that assign preference 1 to two nodes of the variable gadget of x in a non-proper way. We show that any sequence of updates leads to a profile in which all nodes of the gadget adopt preference 0.

Indeed, assume that $w_0(x)$ has preference 1 and both x and $\overline{x}$ have preference 0. Clearly, the nodes $w_1(x)$,..., $w_7(x)$ can switch from 0 to 1 in any order. Among the non-literal nodes $v_i(x)$ and $v_i(\overline{x})$, only one among the degree-3 nodes $v_0(x)$, $v_1(x)$, and $v_1(\overline{x})$ can switch from 0 to 1; this can only happen if the second node with preference 1 is in the neighborhood of one of these nodes (i.e., some of the nodes $v_7(x)$, $v_7(\overline{x})$, $v_2(x)$, or $v_2(\overline{x})$). But then, the literal nodes will have at most four neighbors with preference 1 and they cannot switch to 1. So, no other node has any incentive to switch from 0 to 1. Then, $w_0(x)$ has at least 13 among its 22 neighbors with preference 0 and will switch from 1 to 0, followed by the nodes $w_1(x)$, ..., $w_7(x)$ that will switch back to 0 as well. Then, there are at most two nodes with preference 1 among the nodes $v_i(x)$ and $v_i(\overline{x})$ that will eventually switch to 0 as well (since they have degree at least three).

Assume now that literal node x has preference 1 (the case for $\overline{x}$ is symmetric) and that $w_0(x)$ has preference 0. Then, only the degree-3 node $v_1(x)$ that is adjacent to x can switch to 1 provided that the second node with preference 1 is node $v_2(x)$. Now notice that no other node can switch from 0 to 1. Even worse, the literal node x has at least five (out of nine) neighbors of preference 0 and will switch from 1 to 0. And then, we are left with at most two nodes with preference 1 among the nodes $v_i(x)$ and $v_i(\overline{x})$ that will eventually switch to 0 as well.

Finally, we consider the case in which $w_0(x)$ and the two literal nodes have preference 0. Now the only node that can initially switch from 0 to 1 is $v_0(x)$ provided that the two nodes with preference 1 are $v_7(x)$ and $v_7(\overline{x})$. But then, there is no other node that can switch from 0 to 1 and, eventually, nodes $v_7(x)$ and $v_7(\overline{x})$ will switch to 0 and finally node $v_0(x)$ will switch back to 0.

We have covered all possible cases in which a variable gadget has a non-proper assignment of preference 1 to two nodes and shown that in all of these cases, all nodes of the gadget will switch to preference 0. On the other hand, as discussed above, a proper profile of initial preferences can end up with preference 1 in 17 nodes of the variable gadget.

Now, observe that if at most one node in a clause gadget has preference 1 (or two nodes are assigned preference 1 in a non-proper way), then all the 17 non-clause nodes in the gadget will end up with preference 0. This is due to the fact that none among the nodes $v_1(c)$, ..., $v_{15}(c)$ can switch from 0 to 1 since at least one of their neighbors will have preference 0. But this means that nodes $u_1(c)$ and $u_2(c)$ are adjacent to many (i.e., at least 13) nodes with preference 0; so, they will also switch to 0. And then, if there is still some node $v_i(c)$ with preference 1, it will switch to 0 since both its neighbors have preference 1.

Now, by denoting with V_0, V_1, V_3 the number of variable gadgets that have $0, 1$ or at least 3 nodes with preference 1 and by V_{2p} and V_{2n} the number of variable gadgets with proper and non-proper assignment of preference 1 to exactly two nodes, we have $V = V_0 + V_1 + V_{2n} + V_{2p} + V_3$ and, by denoting with C_0, C_1, and C_3 the number of clause gadgets with 0, 1, and at least 3 nodes with preference 1 in nodes other than the clause node and by C_{2p} and C_{2n} the number of clause gadgets with two nodes with preference 1 assigned in a proper and non-proper way, we have $C = C_0 + C_1 + C_{2n} + C_{2p} + C_3$. Since the total number of nodes with preference 1 does not exceed $2V + 2C$, we have $V_1 + 2V_{2n} + 2V_{2p} + 3V_3 + C_1 + 2C_{2n} + 2C_{2p} + 3C_3 \leq 2V + 2C$ from which we get $V_3 + C_3 \leq 2C_0 + C_1 + 2V_0 + V_1$. Now consider the difference between the number of nodes with preference 1 in any stable profile reached after a sequence of updates and the quantity $17V + 18C$. It is at most $17V_{2p} + 25V_3 + C_0 + C_1 + C_{2n} + 18C_{2p} + 18C_3 - 17V - 18C = -17V_0 - 17V_1 - 17V_{2n} + 8V_3 - 17C_0 - 17C_1 - 17C_{2n} \leq -V_0 - 9V_1 - 17V_{2n} - C_0 - 9C_1 - 17C_{2n} - 8C_3$. Hence, if at least one of V_0, V_1, V_{2n}, C_0, C_1, C_{2n}, and C_3 is positive, the proof follows since the number of nodes with preference 1 will be strictly less than $N + 17V + 18C = n/2$. Otherwise, i.e., if all these quantities are 0, this implies that $C = C_{2p}$ and $V = V_{2p} + V_3$ which in turn implies that $V_3 = 0$ since the number of nodes with preference 1 cannot exceed $2C + 2V$. Hence, the only case where a sequence of updates may lead to a stable profile with at least $N + 17V + 18C$ nodes having preference 1 is when the profile of initial preferences is proper. The claim follows. □

Checking whether minority can become majority. We next show that, given a graph G and a profile of initial preferences $\mathbf{s}_0$, it is possible to decide whether $\mathbf{s}_0$ is mbM for G in polynomial time. Moreover, if this is the case, then there is an efficient algorithm that computes the subverting sequence of updates. This algorithm was used in [5] for bounding the price of stability. Due to page limit, the proof of Theorem 5 is only sketched. We refer the interested reader to the full version of the paper [1].

Theorem 5. *There is a polynomial time algorithm that, given a graph $G = (V, E)$ and a profile of initial preferences $\mathbf{s}_0$, decides whether $\mathbf{s}_0$ is* mbM *for graph G and, if it is, it outputs a subverting sequence of updates.*

Proof (Sketch). Consider the algorithm used in [5] for bounding the price of stability. The running time of the algorithm is polynomial in the size of the input graph, since each node updates its preference at most twice. The fact that the profile $\mathbf{s}'_0$ returned by the algorithm is a stable profile is proved in [5, Lemma 3.3]. We next show that $\mathbf{s}'_0$ is actually the stable profile that maximizes the number of nodes with preference 1. Specifically, consider a sequence σ of updates leading to a stable profile $\mathbf{s}$ that maximizes the number of nodes with preference 1. We will show that there is another sequence of updates that has the form computed by the algorithm described above and converges to a stable profile in which the agents with preference 1 are at least as many as in $\mathbf{s}$. □

4 Conclusions and Open Problems

In this work we showed that, for any social network topology except very few and extreme cases, social pressure can subvert a majority. We proved this with respect to a very natural *majority* dynamics in the case in which agents must express preferences. We also showed that, for each of these graphs, it is possible to compute in polynomial time an initial majority and a sequence of updates that subverts it. The initial majority constructed in this way consists of only $\lceil (n+1)/2 \rceil$ agents. On the other hand, our hardness result proves that it may be hard to compute an initial majority of size at least $3n/4$ that can be subverted by the social pressure. The main problem that this work left open is to close this gap.

Even if computational considerations rule out a simple characterization of the graphs for which a large majority can be subverted, it would be still interesting to gain knowledge on these graphs. Specifically, can we prove that the set of graphs for which large majority can be subverted can be easily described by some simple (but hard to compute) graph-theoretic measure? We believe that our ideas can be adapted (e.g., by considering unbalanced partitions in place of bisections), for gaining useful hints in this direction.

References

1. Auletta, V., Caragiannis, I., Ferraioli, D., Galdi, C., Persiano, G.: Minority becomes majority in social networks. CoRR, abs/1402.4050 (2014)
2. Berger, E.: Dynamic monopolies of constant size. J. Comb. Theory Ser. B **83**(2), 191–200 (2001)
3. Bindel, D., Kleinberg, J.M., Oren, S.: How bad is forming your own opinion? In: Proceedings of the 52nd Annual IEEE Symposium on Foundations of Computer Science (FOCS), pp. 55–66 (2011)
4. Brânzei, S., Caragiannis, I., Morgenstern, J., Procaccia, A.D.: How bad is selfish voting? In: Proceedings of the 27th AAAI Conference on Artificial Intelligence (AAAI), pp. 138–144 (2013)
5. Chierichetti, F., Kleinberg, J.M., Oren, S.: On discrete preferences and coordination. In: Proceedings of the 14th ACM Conference on Electronic Commerce (EC), pp. 233–250 (2013)
6. Coleman, J.S., Katz, E., Menzel, H.: Medical Innovation: A Diffusion Study. Advanced Study in Sociology. Bobbs-Merrill Co., Indianapolis (1966)
7. DeGroot, M.H.: Reaching a consensus. J. Am. Stat. Assoc. **69**(345), 118–121 (1974)
8. Easley, D., Kleinberg, J.: Networks, Crowds, and Markets: Reasoning about A Highly Connected World. Cambridge University Press, New York (2010)
9. Feldman, M., Immorlica, N., Lucier, B., Weinberg, S.M.: Reaching consensus via non-Bayesian asynchronous learning in social networks. In: Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2014), vol. 28, pp. 192–208 (2014)
10. Ferraioli, D., Goldberg, P.W., Ventre, C.: Decentralized dynamics for finite opinion games. In: Serna, M. (ed.) SAGT 2012. LNCS, vol. 7615, pp. 144–155. Springer, Heidelberg (2012)
11. Friedkin, N.E., Johnsen, E.C.: Social influence and opinions. J. Math. Sociol. **15**(3–4), 193–205 (1990)
12. Meir, R., Polukarov, M., Rosenschein, J.S., Jennings, N.R.: Convergence to equilibria in plurality voting. In: Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI), pp. 823–828 (2010)
13. Mossel, E., Neeman, J., Tamuz, O.: Majority dynamics and aggregation of information in social networks. Auton. Agent. Multi-Agent Syst. **28**(3), 408–429 (2014)
14. Mossel, E., Sly, A., Tamuz, O.: Asymptotic learning on Bayesian social networks. Probab. Theory Relat. Fields **158**(1–2), 127–157 (2014)
15. Peleg, D.: Local majorities, coalitions and monopolies in graphs: a review. Theoret. Comput. Sci. **282**, 231–257 (2002)

16. Ryan, B., Gross, N.G.: Acceptance and diffusion of hybrid corn seed in two Iowa communities, vol. 372. Agricultural Experiment Station, Iowa State College of Agriculture and Mechanic Arts (1950)
17. Tamuz, O., Tessler, R.J.: Majority dynamics and the retention of information. Isr. J. Math. **206**(1), 483–507 (2013)
18. Yoshinaka, R.: Higher-order matching in the linear lambda calculus in the absence of constants is NP-complete. In: Giesl, J. (ed.) RTA 2005. LNCS, vol. 3467, pp. 235–249. Springer, Heidelberg (2005)