ECAI 2023
K. Gal et al. (Eds.)

doi:10.3233/FAIA230248

On the Effectiveness of Compact Strategies for Opinion Diffusion in Social Environments

Carlo Adornetto[a], Valeria Fionda[a] and Gianluigi Greco[a]

[a]DeMaCS, University of Calabria, Italy
ORCiD ID: Carlo Adornetto https://orcid.org/0000-0002-9734-1017,
Valeria Fionda https://orcid.org/0000-0002-7018-1324, Gianluigi Greco https://orcid.org/0000-0002-5799-6828

Abstract. An opinion diffusion scenario is considered where two marketers compete to diffuse their own opinions over a social network. In particular, they implement *social proof* marketing approaches that naturally give rise to a strategic setting, where it is crucial to find the appropriate order for targeting the individuals to which provide the incentives to adopt their opinions. The setting is extensively studied from the theoretical and empirical viewpoint, by considering strategies defined in a *compact way*, such as those that can be defined by selecting the individuals according to their degree of centrality in the underlying network. In addition to depicting a clear picture of the complexity issues arising in the setting, several compact strategies are empirically compared on real-world social networks. Results suggest that the effectiveness of compact strategies is moderately influenced by the characteristic of the network, with some centrality measures naturally emerging as good candidates to define heuristic approaches for marketing campaigns.

1 Introduction

The opinions that individuals populating a social environment form and express are significantly impacted by the social pressure of the opinions manifested by their friend/neighbors. Such pressure leads them to exhibit a kind of conformist behaviour, resulting in an *opinion diffusion* process over the underlying network [15, 21]. In fact, the dynamics of opinion diffusion is a central topic of research in areas such as social psychology and political sciences, but it has been recently attracting much attention in the artificial community too (see, e.g., [3, 7, 10, 13, 17, 18, 20] and the references therein).

By abstracting from their specific technical differences, diffusion models can be classified in two main groups [22, 27], namely *progressive* and *non-progressive* ones. In a non-progressive model, an individual that has adopted and manifested an opinion can well change her mind and adopt a different opinion later [25]. Instead, progressive models assume that once an individual adopts an opinion, she remains with that opinion forever. This perspective is appropriate in contexts such as viral marketing [16, 31] or to predict the adoption of new technologies or trends, where the crucial problem is influence maximization [11, 14, 23, 24] via *target set selection*, that is, to identify a small number of individuals that can be profitably used as seeds for a marketing campaign.

In the paper, we precisely consider a progressive scenario in a context where two opinions, say b (black) and w (white), compete for diffusing over the social environment [2, 8, 11, 26, 33]. However, we depart from classical studies related to target set selection, by assuming that the seeds are given and they are not under the control of the marketers, who can instead provide incentives to the individuals to change their opinions in some desired order. In particular, this can practically be done by exhibiting a *social proof* of the opinion, that is, a list of "friends" or influential individuals that have already adopted it. In fact, this setting has been considered in some earlier works in the literature too [4, 5], where it is shown that the specific order used to pick individuals for changing their mind can dramatically affect the number of individuals that eventually hold some desired opinion. However, the questions of how to define an optimal *strategy* for a marketer and of how the strategies of the marketers interplay have been not explored so far, neither from the theoretical viewpoint nor empirically by analyzing the dynamics of some real-world networks.

Our work embarks in a systematic study of the above questions within a setting where marketing strategies are defined in a *compact way*. Indeed, in a general setting, a diffusion strategy might well depend on the history of the evolution of the network as well as by the specific configuration given to hand. However, modeling and reasoning with such arbitrary strategies would require extremely demanding computational resources and an assumption of complete knowledge, which is unrealistic in real-world scenarios. In fact, in order to identify the individuals to target for the propagation of the opinion, marketers are often guided by some heuristic parameters aimed at estimating the social "power" of the individuals in the network. A noticeable example is when such power is modeled in terms of some well-known *centrality measures* [30], and where a strategy might be specified by just picking the individual with the highest rank (according to the desired measure) over all possible individuals that can potentially change their mind with a social proof marketing strategy.

In more details, we provide the following contribution:

- ► We define a strategic setting for reasoning about progressive dynamics determined by compact strategies. Our modeling takes care of the *speed* of the propagation, which reflects the efforts spent by the marketer to spread her opinion, and of the interplay between the strategies of the competing opinions. Moreover, to define the social pressure of the individuals, we assume a deterministic *linear threshold* [22] setting, that is, an individual can adopt an opinion only if (at least) a given fraction of her neighbors did.
- ► We formalize some relevant problems arising in our strategic setting and we study their computational complexity. The study is conducted for strategies defined in a compact way, but not necessarily restricted to those induced by centrality measures.

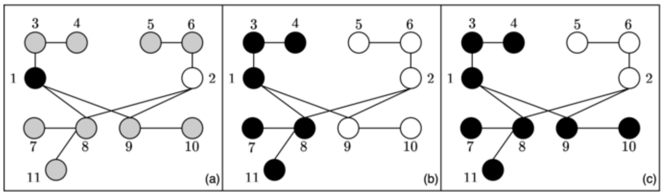

Figure 1. Illustrations for examples in Section 2.

- And, finally, we focus on some well-known centrality measures and we performed an extensive experimental evaluation aimed at assessing the quality of the strategies they naturally induce. Our campaign considers real social networks, and diffusion processes with different characteristics determined by the initial seeds and the speed of propagation.

The rest of the paper is organized as follows. The model for opinion diffusion is presented in Section 2. Our formal and empirical studies are reported in Section 3 and Section 4, respectively. Some final remarks on our findings as well as some directions for further works are eventually discussed in Section 5.

2 Compact Strategies for Opinion Diffusion

Social networks and dynamics. A social network is modeled as an undirected graph $G = (N, E)$ over a set N of individuals/nodes. Two competing opinions, denoted as b (*black*) and w (*white*), are spreading over the network, by starting from some initial seeds. A special opinion g (*gray*) is associated with any individual $v \in N$ that has not already adopted an opinion in $\{\mathsf{w}, \mathsf{b}\}$; in this case, v is influenced by her *neighbors* in G, i.e., by the individuals in the set $\delta(v) = \{x \mid \{v, x\} \in E\}$, under a *linear threshold model* [22] for some fixed threshold $0 \leq t \leq 1$.

Formally, a *configuration* for G is defined as a pair $S = (S_\mathsf{b}, S_\mathsf{w})$ such that $S_\mathsf{b}, S_\mathsf{w} \subseteq N$ is the set of all individuals that hold opinion b and w, respectively. For each individual $v \in N \setminus (S_\mathsf{b} \cup S_\mathsf{w})$ that has not already adopted an opinion, let $\sigma(v) = \lceil t \cdot |\delta(v)| \rceil$. Then, we say that v is *stable* with respect to the configuration $(S_\mathsf{b}, S_\mathsf{w})$ if $|\delta(v) \cap S_\mathsf{b}| < \sigma(v)$ and $|\delta(v) \cap S_\mathsf{w}| < \sigma(v)$. The configuration S is stable if all individuals in $N \setminus (S_\mathsf{b} \cup S_\mathsf{w})$ are stable. A *dynamic* for G is a sequence of configurations $\pi = (S^0, ..., S^k)$ such that S^k is stable and, for each $i \in \{1, ..., k\}$, S^i is obtained from S^{i-1} by picking an individual in $N \setminus (S^{i-1}_\mathsf{b} \cup S^{i-1}_\mathsf{w})$ that is not stable in S^{i-1} and by setting her opinion to b or w. Note that we are considering *progressive* dynamics; therefore, for each *initial configuration* S^0, $k \leq |N \setminus (S^0_\mathsf{b} \cup S^0_\mathsf{w})|$ always holds.

Example 1 Consider the network in Figure 1(a), the configuration $(\{1\}, \{2\})$ and the threshold $t = 0$ (meaning that one neighbor that is not g is enough to make the node not stable). Then, the configuration is not stable, as individuals in $\{3, 8, 9\}$ can change their opinion to b, and individuals in $\{6, 8, 9\}$ can change to w. $\lhd$

Compact Startegies. We consider a strategic setting for opinion diffusion, where dynamics originate from the interactions of two players, say P_b and P_w, competing to maximize the spread of b and w, respectively. Formally, a strategy for P_b (resp., P_w) is a function τ_b (resp., τ_w) associating with any configuration S over the network G a node $\tau_\mathsf{b}(S) \in N$ (resp., $\tau_\mathsf{w}(S) \in N$) that is not stable and that can change her opinion to b (resp., w). In particular, note that we are considering strategies that do not depend on the *history* of the evolution of the network, which is a rather natural assumption in all those settings where strategies are a-priori defined in terms of structural/topological properties of the social network. For instance, a strategy of interest to our analysis can be the one of selecting the individual that can change the opinion and that have the maximum possible degree. Such strategies will be hereinafter called *compact*. In formal terms, a strategy τ_b (resp., τ_w) is compact if it is given as a polynomial-time computable function defined over some internal encoding, say $\epsilon(\tau_\mathsf{b})$ (resp. $\epsilon(\tau_\mathsf{w})$), whose size is polynomially bounded in the size of G. In fact, if a strategy is not compact, than its encoding would naturally require to list all possible network configurations with their associated outcomes, hence requiring exponential space (rather than polynomial). We refer the reader to Section 4 for further relevant compact strategies that we consider in our experimentation.

Speed of Diffusion. We assume that players act in turns. Moreover, as a way to formalize the efforts spent in spreading their opinions, we define the *speed* of diffusion as a pair $\rho = (\rho_\mathsf{b}, \rho_\mathsf{w})$ of natural numbers characterizing, at each turn, the number of individuals selected by P_b or by P_w, respectively, to change their mind.

In fact, given an initial configuration S^0 and the speed ρ, the strategies τ_b and τ_w univocally determine a dynamic $S^0, ..., S^k$ for G, which we hereinafter denote as $\pi[S^0, \rho, \tau_\mathsf{b}, \tau_\mathsf{w}]$ and which is defined as follows. W.l.o.g., the first turn of player P_b starts in S^0. When the turn of P_b (resp., P_w) starts in some configuration S^i, then the dynamic evolves by iteratively changing the mind to m individuals according to τ_b (resp., τ_w), such that either $m = \rho_\mathsf{b}$ (resp., $m = \rho_\mathsf{w}$) or S^{i+m} contains no individual that can change her opinion to b (resp., w); eventually, the turn of the other player starts in S^{i+m+1}.

Example 2 Consider a strategy D_b (resp., D_w) for P_b (resp., P_w) that selects, for each configuration S, the node that is not stable in S and can change her opinion to b (resp., w) having the maximum degree. By starting from the configuration in Figure 1(a), and by considering the speed $(1, 1)$, the network evolves as follows: $8 \mapsto \mathsf{b}; 9 \mapsto \mathsf{w}; 3 \mapsto \mathsf{b}; 6 \mapsto \mathsf{w}; 4 \mapsto \mathsf{b}; 5 \mapsto \mathsf{w}; 7 \mapsto \mathsf{b}; 10 \mapsto \mathsf{w}; 11 \mapsto \mathsf{b}$. Eventually, the dynamic induced by D_b and D_w, say $\dot{\pi}$, will lead to the stable configuration reported in Figure 1(b). $\lhd$

Coverage. In the following, we shall study opinion diffusion from the perspective of maximizing the spread of the opinions b and w. Hence, in order to finalize the formalization of the framework, it is natural to define the *coverage* of b (resp, w) over G of the given dynamic $\pi = S^0, ..., S^k$ as the number $\gamma_\mathsf{b}(\pi)$ (resp., $\gamma_\mathsf{w}(\pi)$) of individuals holding opinion b (resp., opinion w) at the end of π, that is $\gamma_\mathsf{b}(\pi) = |S^k_\mathsf{b}|$ (resp., $\gamma_\mathsf{w}(\pi) = |S^k_\mathsf{w}|$).

Example 3 The coverage of b (resp., w) in the dynamic $\dot{\pi}$ of Example 2 is $\gamma_\mathsf{b}(\dot{\pi}) = 6$ (resp., $\gamma_\mathsf{w}(\dot{\pi}) = 5$). $\lhd$

3 Reasoning about Opinion Diffusion

Now that we have defined a formal framework for reasoning about opinion diffusion under compact strategies, we can turn to study some relevant computational problems arising therein. In particular, we next embark on the definition and study of the opinion maximization problem by considering two kinds of setting determined by the strategic interplay emerging between players P_b and P_w.

3.1 Brave and Cautious Reasoning

Let S^0 be an initial configuration and ρ be the speed. Recall that we are considering strategies τ_{b} and τ_{w} that are functions of the configuration S at hand only. We take the perspective of player P_{b} and we assume that τ_{w} is private to w. In particular, P_{b} is in charge of determining her best possible strategy and, given the uncertainty about τ_w, two approaches can be considered.

- On the one hand, player P_{b} might take an optimistic perspective, according to which her coverage is defined as the maximum possible coverage over all the possible strategies of τ_{w}. Accordingly, we say that the strategy $\tau_{\mathsf{b}}^{\top}$ is *brave*-optimal for P_{b} if there exists a strategy $\tau_{\mathsf{w}}^{\top}$ for P_{w} such that:

$$(\tau_{\mathsf{b}}^{\top}, \tau_{\mathsf{w}}^{\top}) = \arg\max_{\tau_{\mathsf{b}}, \tau_w} \gamma_{\mathsf{b}}(\pi[S^0, \rho, \tau_{\mathsf{b}}, \tau_{\mathsf{w}}]).$$

- On the other hand, player P_{b} might take a pessimistic viewpoint, in that she assumes that P_{w} always plays the strategy that maximally reduce her coverage. Accordingly, we say that the strategy $\tau_{\mathsf{b}}^{\perp}$ is *cautious*-optimal for P_{b} if:

$$\tau_{\mathsf{b}}^{\perp} = \arg\max_{\tau_b} \left(\min_{\tau_w} \gamma_{\mathsf{b}}(\pi[S^0, \rho, \tau_{\mathsf{b}}, \tau_{\mathsf{w}}]) \right).$$

Example 4 Consider the network and the initial configuration reported in Figure 1(a). Assume that P_{b} adopts the strategy D_{b} of selecting, for each configuration S, the individual having the maximum degree that is not stable in S and can change her opinion to b. Then, according to a cautious perspective, the maximum coverage that can be obtained is the one in Figure 1(b) and discussed in Example 2.

Consider now the brave perspective. In this case, the maximum coverage for P_{b} is associated with the strategy M_{w} for P_{w} that selects, for each configuration S, the node that is not stable in S and can change her opinion to w, having the minimum degree. By starting from the configuration in Figure 1(a), and by considering the speed $(1, 1)$, the network evolves as follows: $8 \mapsto \mathsf{b}; 6 \mapsto \mathsf{w}; 9 \mapsto \mathsf{b}; 5 \mapsto \mathsf{w}; 3 \mapsto \mathsf{b}; 4 \mapsto \mathsf{b}; 7 \mapsto \mathsf{b}; 10 \mapsto \mathsf{b}; 11 \mapsto \mathsf{b}$. Eventually, the dynamic $\overline{\pi}$ induced by D_{b} and M_{w}, will lead to the stable configuration reported in Figure 1(c) where the coverage of b is $\gamma_{\mathsf{b}}(\overline{\pi}) = 8$. $\lhd$

Armed with the above notions, we can naturally define the following two (Opinion Maximization) problems, receiving as input G, the initial configuration S^0, the speed ρ, and a real number $\alpha \in [0, 1]$:

BRAVE-OM: Is $\gamma^b(\pi[S^0, \rho, \tau_{\mathsf{b}}^{\perp}, \tau_{\mathsf{w}}^{\perp}]) \geq \alpha \times |N|$?
CAUTIOUS-OM: Is $\gamma^b(\pi[S^0, \rho, \tau_{\mathsf{b}}^{\perp}, \tau_{\mathsf{w}}]) \geq \alpha \times |N|$, for each possible strategy τ_{w} for P_{w}?

The complexity of these two problems will be next analyzed.

3.2 Complexity Analysis

Given that we are considering compact strategies, it is immediate to check that problem BRAVE-OM belongs to the class NP of all problems that can be solved in polynomial time by a non-deterministic Turing machine. Indeed, we can just guess the strategy $\tau_{\mathsf{b}}^{\perp}$ (whose encoding $\epsilon(\tau_{\mathsf{b}}^{\perp})$ requires polynomially-many bits) and then check in polynomial time whether $\gamma^b(\pi[S^0, \rho, \tau_{\mathsf{b}}^{\perp}, \tau_{\mathsf{w}}^{\perp}]) \geq \alpha \times |N|$ actually holds. Things are more complex with CAUTIOUS-OM. Indeed, in this case, we can still guess in polynomial time $\tau_{\mathsf{b}}^{\top}$ over a non-deterministic Turing machine; but, now the problem of checking whether $\gamma^b(\pi[S^0, \rho, \tau_{\mathsf{b}}^{\perp}, \tau_{\mathsf{w}}]) \geq \alpha \times |N|$ holds for each τ_w requires solving another problem in NP (which amounts at checking whether there exists some τ_{w}^* such that $\gamma^b(\pi[S^0, \rho, \tau_{\mathsf{b}}^{\perp}, \tau_{\mathsf{w}}^*]) < \alpha \times |N|$). Hence, CAUTIOUS-OM belongs to the complexity class Σ_2^{P} [28].

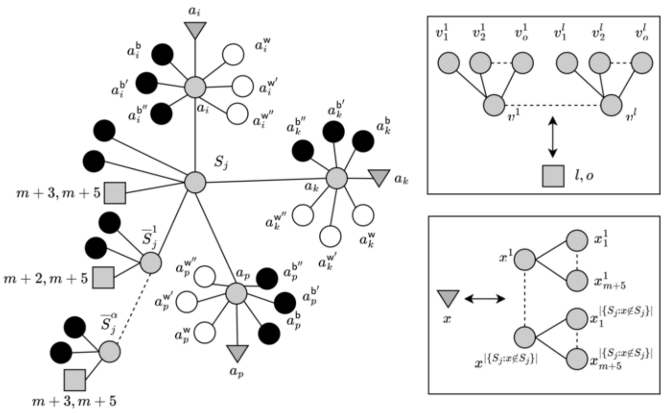

Figure 2. Illustration of the reduction in the proof of Theorem 5.

We next complete the picture by showing the above results are tight. In fact, we start by showing that BRAVE-OM is NP-hard, by exibithing a reduction to the 3Hitting Set problem [19] (shortly 3HS), that, given a collection $C = \{S_1, \ldots, S_m\}$ of subsets of size three of a finite set $S = \{a_1, ..., a_n\}$ and an integer k, is the problem of checking whether there exists a subset $S' \subseteq S$ such that $|S'| \leq k$ and S' contains at least one element from each subset in C.

Theorem 5 BRAVE-OM *is* NP-*complete.*

Proof (Sketch). Let $C = \{S_1, ..., S_m\}$ be a collection of subsets of size three of a finite set $S = \{a_1, ..., a_n\}$ and let k be a positive integer. Consider the network $G = (N, E)$, depicted in Figure 2, where we have one node a_i for each element $a_i \in S$ and a node S_j for each subset $S_j \in C$ – in the figure nodes a_i, a_k, a_p represent the elements belonging to the set S_j. Note that the network is built such that each node has exactly 1 or $m+6$ neighbors, and from each S_j node there is a chain of g nodes $\overline{S}_j^1, ..., \overline{S}_j^\alpha$ of length α, where α is chosen to be far greater than $\max\{m, n\}$. Consider a threshold $t = 3/(m+6)$, then, according to such a threshold, only a_i, S_j and $\overline{S}_j^t$ nodes can change their opinions, since the gadgets reported in the right part of Figure 2 prevent other nodes of being able to change their opinion. We now claim that the (C, S, k) is a *yes* instance of 3HS if, and only if, BRAVE-OM returns *yes* on G, with the initial configuration reported in Figure 2, where all nodes are $3/(m+6)$-individuals, and by considering a speed of diffusion $(k, n-k)$ and a final coverage threshold of $3 * m * \alpha/|N|$.

(if part) Let S' be a set witnessing that (C, S, k) is a *yes* instance to 3HS. A strategy for b is to diffuse to the nodes $a_i \in S'$ in the first k steps (if $|S'| < k$, the remaining $k - |S'|$ nodes can be chosen randomly among the remaining a_i). Then, the only possibility for w in the subsequent $n - k$ steps is to diffuse in the a_i that are still g. Then, w cannot diffuse anymore. In fact, note that all three a_i, a_k, a_p must be w to enable S_j to switch to w, while just one of them is required to be b for S_j being able to switch to b. Since S' is a solution to 3HS, it means that at least one element, say a_i, for each $S_j \in C$ is in S' and, thus, the corresponding node a_i has switched to b in the first k steps of the dynamic. Thus, in the subsequent steps all nodes S_j will switch to b thus enabling the $m\,\overline{S}_j$ chains of length α to also switch to b. This conclude the proof since at the end of the dynamic the number of nodes holding opinion b is greater than $3 * m * \alpha$.

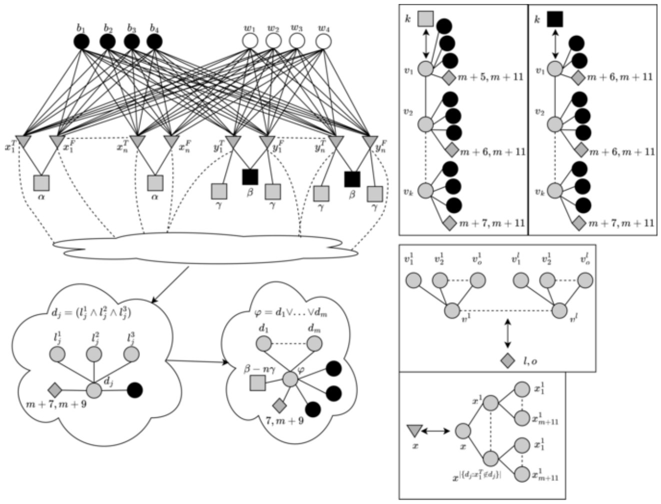

Figure 3. Illustration of the reduction in the proof of Theorem 6.

(only-if part) Let $(S_k^{\mathsf{b}}, S_k^{\mathsf{w}})$ be the final configuration of a dynamic witnessing that BRAVE-OM returns *yes*. Note that, to be able to reach the coverage threshold, all S_j nodes had to become b at some point so enabling the m $\overline{S}_j$ chains of length α to switch to b too. Thus, it means that for each node S_j there is at least one a_i neighbor holding opinion b. The set $S' = \{a_i \mid a_i \in S_k^{\mathsf{b}}\}$ covers all the subsets $S_j \in C$. To conclude, note that according to the speed of diffusion the nodes a_i that have opinion b are at most k (and these are the nodes that changed their opinion in the first k steps), since opinion w can only cover the other a_i nodes, that are (at least) $n-k$, since they are the only nodes that are enabled to switch to w in every dynamic witnessing a *yes* instance of BRAVE-OM. □

We next complete the picture by showing that CAUTIOUS-OM is Σ_2^{P}-hard, by exhibiting a (rather elaborated) reduction to the problem of deciding the validity of a Quantified Boolean Formula (QBF) having the form $\exists x \forall y.\varphi$ and where φ is given in disjunctive normal form [28]. Practically, this means that—unlike BRAVE-OM, that is a "classical" NP-complete problem–we cannot design a *flat-backtracking* algorithm for CAUTIOUS-OM, i.e., where the search-space is a tree having a polynomial number of levels (and such that moving along the tree edges does not take exponential time).

Theorem 6 CAUTIOUS-OM *is* Σ_2^{P}*-complete.*

Proof (Sketch). Let $\exists x_1, ..., x_n \forall y_1, ..., y_n \varphi$ be a 2QBF formula in disjunctive normal form with m clauses[1]. Consider the network $G = (N, E)$, depicted in Figure 3, where we have two nodes v_i^T and v_i^F for each variable v_i, that are meant to encode the truth assignment for variable v_i. Moreover, there is a node d_j for each disjunct (in the figure nodes l_j^1, l_j^2, l_j^3 represent the nodes associated to the variables directed or negated that appear in the disjunct d_j) and a node φ representing the formula. Note that the network is built such that each node has exactly 1 or $m+11$ neighbors (see the gadgets reported in the right part of the figure), and from each x_i^T and x_i^F node there is a chain of g nodes of length α, while from each y_i^T and y_i^F node there are a chain of g nodes of length γ and a chain of g nodes of length β with $\alpha \gg \beta \gg \gamma \gg \max(n, m)$. Furthermore, there is also a chain of g nodes of length $\beta - n\gamma$ starting from the node φ. If we consider a threshold $t = 4/(m+11)$, only the nodes x_i^T, x_i^F, y_i^T, y_i^F, d_j and φ (as well as the α, β and γ chains) can change their opinions, since the gadgets reported in the right part of Figure 3 prevent other nodes of being able to change their opinion.

We now claim that $\exists x_1, ..., x_n \forall y_1, ..., y_n \varphi$ is valid if, and only if, the answer to CAUTIOUS-OM is *yes* on G, with the initial configuration reported in Figure 3, where all nodes are $4/(m+11)$-individuals, and by considering a speed of diffusion (n, n) and a final coverage threshold of $(n * \alpha + \beta)/|N|$.

(if part) Let $\mathbb{X}$ be a satisfying assignment for the existentially quantified variables witnessing the validity of $\exists x_1, ..., x_n \forall y_1, ..., y_n \varphi$. A strategy for b is to diffuse to nodes x_i^T (resp., x_i^F) for each x_i that evaluates true (resp., false) in $\mathbb{X}$ in the first n steps. Then, w can diffuse to the x nodes that are still g in the subsequent n steps. From this point, we can consider whatever truth assignment for variables y, thus we can assume that b will diffuse to $y_1^T, ..., y_n^T$ and w to $y_1^F, ..., y_n^F$. From this configuration only b is enabled to diffuse to the n chains of length α connected to the x nodes and to the n chains of length γ connected to the y_i^T nodes. Moreover, since $\mathbb{X}$ is witnessing the validity of the formula, it means that there is at least one disjunct, say d_j, that evaluates true in $\mathbb{X}$. This means that the three nodes l_j^1, l_j^2, l_j^3 associated to the literals that appear in d_j hold opinion b and enable node d_j to adopt opinion b too. To conclude, note that after d_j adopts opinion b also node φ becomes not stable and can adopt opinion b, thus enabling the last chain of length $\beta - n\gamma$ to change its opinion to b.

(only-if part) Let $(S_{\mathsf{b}}^k, S_{\mathsf{w}}^k)$ be the final configuration of a dynamic π witnessing that the answer to CAUTIOUS-OM is *yes*, and consider the truth assignment $\mathbb{X}$ such that x_i evaluates true (resp., false) in $\mathbb{X}$ if x_i^T (resp., x_i^F) becomes b in the first n steps of π. By definition of π, we have that $|S_{\mathsf{b}}^k| \geq n * \alpha + \beta$. Note that, to meet such a requirement it is mandatory that all α chains connected to the x nodes must be b in S_{b}^k and thus, for each x_i at least one among x_i^T and x_i^F must be b. Moreover, according to the CAUTIOUS-OM setting, the strategy selected from b must allow to obtain a valid solution for whatever strategy adopted by w, and thus b must diffuse in the first n steps to exactly one node between x_i^T and x_i^F for each x_i. In fact, suppose that b diffuse to both x_i^T and x_i^F for some x_i, it means that there exists an x_j for which both x_j^T and x_j^F are still g after the first n steps and to which w can diffuse by preventing b to subsequent diffuse in the corresponding α chain. Then, we will show that whatever strategy played by w is always a winning strategy for b.

If w in the subsequent n steps will leave free two nodes y_i^T and y_i^F for some y_i, then b can diffuse to both of them thus enabling the corresponding β chain to become b and meeting the coverage requirement. If, on the contrary, w diffuses to either y_i^T or y_i^F for each y_i (thus, for each valid truth assignment), then b can diffus e to the y nodes that are still g in the subsequent n steps, enabling n chains of length γ to adopt opinion b too[2]. Then, to meet coverage requirement there are still $(\beta - n\gamma)$ nodes missing that can obtained only via the $(\beta - n\gamma)$ chain connected to the φ node. To enable such a chain to switch to b, node φ must switch to b too, meaning that at least one d_j is enabled to switch to b because its three literals nodes l_j^1, l_j^2, l_j^3 became b in the previous steps of the dynamics. □

[1] Note that, w.l.o.g., we consider a formula with $2n$ variables, n quantified existentially and n quantified universally.

[2] Note that, if b decides to diffuse in some x_i^T/x_i^F still g then w can occupy an y_p^T/y_p^F left g thus preventing b to obtain a γ chain and, thus, it is not a valid strategy for b.

Network	$\lvert N\rvert$	$\lvert E\rvert$	r_{kk}	$\langle k\rangle$	k^*	λ
dblp	317080	1049866	0.27	6.6	343	1.5
fb	134873	1380293	0.07	20.5	1469	1.3
deezer	143884	846915	0.33	11.8	420	1.3
fb-Art	50521	819306	-0.02	32.4	1469	1.2
fb-Ath	13868	86858	-0.03	12.5	468	1.3
fb-Com	14120	52310	0.01	7.4	215	1.4
fb-Gov	7058	89455	0.03	25.3	697	1.2
fb-NS	27930	206259	0.02	14.8	678	1.3
fb-Pol	5908	41729	0.02	14.1	323	1.3
fb-PF	11573	67114	0.20	11.6	326	1.4
fb-TvS	3895	17262	0.56	8.9	126	1.3
dz-HR	54573	498202	0.20	18.3	420	1.2
dz-HU	47538	222887	0.21	9.4	112	1.2
dz-RO	41773	125826	0.11	6.0	112	1.4

Figure 4. Dataset characteristics for the experiments in Section 4.

4 Compact Strategies via Centrality Measures

In this section, we turn to study opinion diffusion from an empirical viewpoint by focusing on an important class of compact strategies, namely those that are naturally identified by the ranking induced by a centrality measure [30]. In fact, each measure naturally induces a strategy where the next individual to be picked is the one with the highest rank over the individuals that can adopt the given opinion.

In particular, in our analysis, we shall consider the following centrality measures, which—without ambiguity—will be hereinafter transparently referred to as strategies:

- *degree centrality* (deg), which counts the number of connections of an individual;
- *betweenness centrality* (bet), which measures the number of shortest paths that pass through a particular individual;
- *closeness centrality* (cls), which looks at the average distance from a particular individual to all others in the network;
- *eigenvector centrality* (eig), which considers both the number of connections that an individual has, as well as the centrality of the individuals that it is connected to;
- *vote rank* (vr), which is based on the concept of voting, where each individual in the network has the ability to cast a vote for other individuals; the more votes a node receives, the higher its vote rank centrality measure.

The quality of these measures/startegies (from the perspective of opinion diffusion) is next assessed by considering a thorough experimentation conducted on several real-world social environments.

4.1 Experimental Setting

Dataset. To test the effectiveness of different compact strategies on opinion diffusion we used several real networks. We considered a benchmark consisting of 13 graph datasets, whose main features are summarized in Figure 4. In particular, for each dataset, we report the number of nodes $|N|$, the number of edges $|E|$, the assortativity r_{kk}, the average node degree $\langle k\rangle$, the maximum node degree k^* and the coefficient of an (approximate) underlying power law distribution λ. In more detauils, assortativity (or degree correlation) r_{kk} is the Pearson correlation between the degrees of connected nodes. In assortative networks ($r_{kk} > 0$) nodes are connected to nodes having similar degree, while in disassortative networks ($r_{kk} < 0$) they link

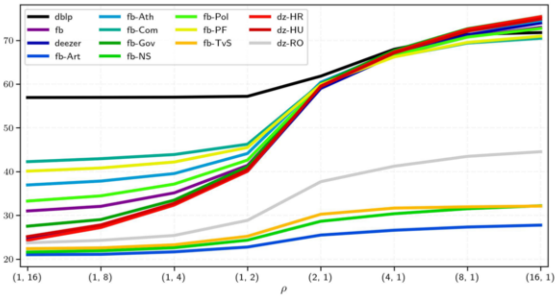

Figure 5. Final percentage of individuals holding opinion b according to different speed of diffusion (ρ_{b},ρ_{w}), when the strategy for both opinions is deg, the threshold is 0.1 and initial seed ratios for both opinions is 0.2.

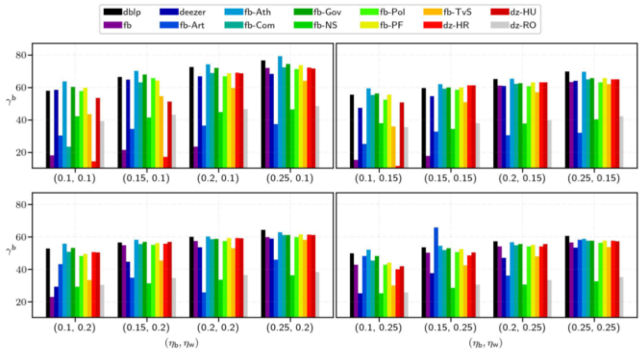

Figure 6. Final percentage of individuals holding opinion b according to different ratios (η_{b},η_{w}) of nodes in the initial configuration, when the strategy for both opinions is deg, the threshold is 0.3, and the speed of diffusion is $(2, 1)$. Each subplot is obtained by fixing η_{w}, and varying η_{b}.

to nodes having dissimilar degree. Note that two networks having the same degree distribution can differ for their assortativity. Moreover, the coefficient of the power law distribution (i.e., λ) that better approximate the degree distribution of the network has been computed according to the Bhattacharyya distance [6].

The datasets fb-Art, fb-Ath, fb-Com, fb-Gov, fb-NS, fb-Pol, fb-PF, fb-TvS have been extracted from the Facebook (fb) [29] dataset by considering artists, athletes, companies, governments, new sites, politicians, public figures and TV shows pages only, respectively. The datasets dz-HR, dz-HU and dz-RO have been extracted from the Deezer dataset [29] by considering the friendships networks of users in Croatia, Hungary and Romania, respectively. The largest dataset is dblp [32] having more than 300K nodes and 1M edges.

Experimental Setup. Experiments have been conducted as follows. For each pair of strategies $\tau_{\mathsf{b}}, \tau_{\mathsf{w}} \in \{\texttt{deg}, \texttt{bet}, \texttt{cls}, \texttt{eig}, \texttt{vr}\}$, we varied the number of individuals having opinions b and w in the initial configuration. In particular, for each pair of real numbers $\eta_{\mathsf{b}}, \eta_{\mathsf{w}} \in \{0.1, 0.15, 0.20, 0.25\}$, the initial configuration $(S^0_{\mathsf{b}}, S^0_{\mathsf{w}})$ was determined as follows: S^0_{b} consists of the $\eta_{\mathsf{b}} * |N|$ nodes having the highest value according to τ_{b}, while S^0_{w} consists of the $\eta_{\mathsf{w}} * |N|$ nodes having the highest value according to τ_{w} in $N \setminus S^0_{\mathsf{b}}$ (individuals still holding opinion g after b initialization). Then, for each pair of strategies and each initial configuration, we considered different values for the threshold t determining when a node is not stable, by considering $t \in \{0.1, 0.3, 0.35, 0.4, 0.45, 0.5\}$. Finally, we also varied the relative speed of diffusion of b and w by considering $(\rho_{\mathsf{b}}, \rho_{\mathsf{w}}) \in \{(1, 16), (1, 8), (1, 4), (1, 2), (2, 1), (4, 1), (8, 1), (16, 1)\}$.

Overall, for each network in Figure 4, we considered 19.200 different experimental settings from which we simulated the diffusion

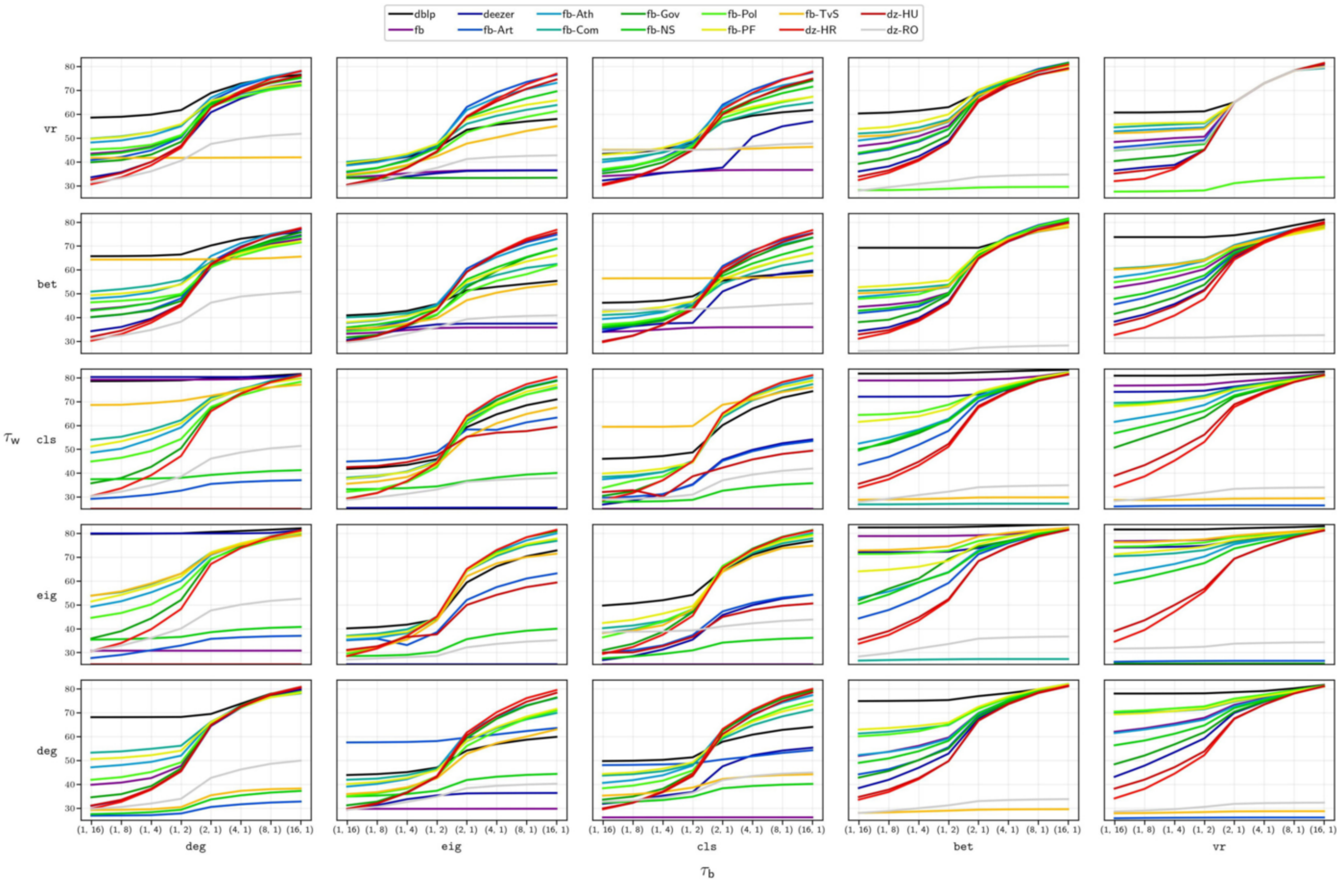

Figure 7. Final percentage of individuals holding opinion b for each pair of strategies for b and w (on outer x-axis and y-axis respectively), according to different speed (on the subplots x-axis), for $t = 0.1$, $\eta_b = 0.25$ and $\eta_w = 0.15$.

dynamics according to the strategies and the speed given at hand.

Execution environment. All experiments were executed in an Anaconda3 virtual environment, where an opinion diffusion framework has been implemented in Python (v3.8.13) by taking advantage of the NetworkX (v2.8.6) library. All experiments ran on high-performance computing node with Intel(R) Xeon(R) Gold 5118 CPU (2.30GHz), 4x12 cores double thread (for a total of 96 threads) and 512GB of RAM. In order to take full advantage of the available resources, the experiments were fully parallelized such that 96 different settings for a graph ran simultaneously on separate CPU threads.

4.2 Results

In order to shed lights on the behaviour of the various compact strategies, we start by discussing the results of our experimental campaign by looking at how the final coverage of opinion b is related to the different parameters that determine the experimental settings. Figure 5 reports the final coverage of b (in terms of percentage of nodes holding opinion b at the end of the diffusion process) for increasing values of the relative speed of diffusion for b; in particular, we depict the results for $\tau_b = \tau_w = \texttt{deg}$, $\eta_b = \eta_w = 0.2$, and $t = 0.1$. Our findings—which are representative of the behaviour manifested over all the other settings—evidence that higher speeds lead to an increase of the coverage. Moreover, note that a steeper increase in the coverage emerges as soon as ρ_b becomes greater than ρ_w. Figure 6 shows, instead, the coverage of b, when both opinions adopt the $\texttt{deg}$ strategy and diffuse with a relative speed of $(2, 1)$, by considering the threshold 0.3, w.r.t. different initial configurations. In particular, each sub-chart considers η_b as a variable for a fixed η_w. As can it be noted, an increase in the number of initial seeds does not necessarily correspond to an increase of the final coverage. In fact, when η_w is 0.1 (upper-left chart) and 0.15 (upper-right chart) the final coverage of b in the fb-NS network decreases when η_b increases from 0.1 to 0.15. Similarly, also for the graph fb-Art for all η_w except 0.1, the final coverage of b decreases when η_b increases from 0.15 to 0.2.

A clear pictures of the effectiveness of the various strategies considered in the experimentation is then reported in Figure 7. In particular, we report the final coverage of b (again, as percentage of the whole network) for different strategies for b (outer x-axis) and w (outer y-axis), and speed of diffusion (x-axis of subplots)—results are referred to a threshold $t = 0.1$ and to the ratios of b and w in the initial configuration equals to 0.25 and 0.15, respectively. As it can be noted, the impact of the strategy τ_b on the final coverage heavily depends on the strategy τ_w. For example, on the Facebook network (fb), when considering $\tau_b = \texttt{deg}$ and the speed of diffusion $(16, 1)$, the final coverage of b approaches the 80% of nodes if $\tau_w \in \{\texttt{deg}, \texttt{cls}, \texttt{bet}, \texttt{vr}\}$ but is around the 30% of nodes when $\tau_w = \texttt{eig}$. A similar behaviour can be noted on the deezer network when considering $\tau_b = \texttt{eig}$. In fact, for the speed $(16, 1)$ the final coverage of b is around 0% of nodes if $\tau_w \in \{\texttt{eig}, \texttt{cls}\}$ but reaches the $35\% - 40\%$ of nodes when $\tau_w \in \{\texttt{deg}, \texttt{bet}, \texttt{vr}\}$. Finally, by

Figure 8. Total number of settings in which `vr` and `bet` resulted to be brave- or cautious-optimal.

looking at Figure 7, we note that the efficacy of some strategies depends on the characteristics of the network. For example, `deg` and `bet` are much less effective on dz-RO than on dblp, while `cls` on dz-RO is more effective than on dblp in several settings.

Finally, to sum up all the findings of our experimental campaign, we depict a synthetic picture in Figure 8 which can be used to identify the best strategies according to both brave and cautious reasoning. Indeed, Figure 8 reports the overall number of settings (over all input networks and by considering all initial configurations) in which each strategy has been identified as brave-optimal and cautious-optimal (for P_{b}) over the total of 1792 settings for each threshold. In Figure 8, subplots shows the results for different thresholds. In each subplot, the first five rows consider settings where the speed of diffusion of b is lower than the speed of diffusion of w (i.e., $(1, 2)$, $(1, 4)$, $(1, 8)$ and $(1, 16)$) while rows $6 - 10$ consider settings in which the speed of b is greater than the speed of w (i.e., $(2, 1)$, $(4, 1)$, $(8, 1)$ and $(16, 1)$). For example, by considering $t = 0.1$, `bet` was cautious-optimal in 938 settings, while for $t = 0.45$ it resulted to be cautious-optimal in 503 settings only. From the analysis we performed, `bet` seems to be the best *cautious* choice for low thresholds, while in all the other cases the best cautious choice is `vr`. As for brave reasoning, each cell reports the number of settings in which a strategy has been identified as brave-optimal together with the corresponding strategy for w. For example, by considering $t = 0.4$ and $\rho_{\mathsf{b}} > \rho_{\mathsf{w}}$, `vr` for b was brave-optimal in 569 settings together with `eig` for w (that resulted in this settings the worst strategy for w), while `bet` for b was brave optimal in 28 settings together with `cls` for w. From the analysis of the results on the one hand it emerged that, in general, according to a brave-reasoning, the worst strategy for w id `eig`, followed by `deg`, that are those allowing b to reach the maximum coverage for some strategy τ_{b}. On the other hand, again `vr` is the brave-optimal strategy for b in the majority of the settings, followed by `bet`.

5 Discussion and Conclusion

Opinion diffusion has been largely studied in earlier literature. Several studies considered a setting in which there are two opinions that compete [2, 8, 11], and some recent works also considered the scenario in which more than two opinions are available [1, 5, 9, 12]. In this paper, we have analyzed a progressive model of opinion diffusion, in which individuals can hold one of two competing opinions or can have no opinion at all. Given this model, we have investigated the effectiveness of *compact* strategies that, if adopted by marketers, suggests the sequence of individuals to target for maximizing the final diffusion of their opinion. We studied this problem from a theoretical point of view and complemented our analysis with an experimental evaluation that demonstrate how compact strategies can be effectively adopted in opinion maximization. In particular, our findings suggest that *vote rank* and *betweenness centrality* are very effective measure to characterize the power of the nodes in terms of their capacity to affect the final coverage of the diffusion process.

Our results open a number of avenues for further research. First, it would be relevant to investigate the impact of further network characteristics, such as network density, on the effectiveness of compact strategies. Furthermore, while we conducted experiments on large real social networks, it might be nonetheless interesting to consider small synthetic networks on which it would be possible to check how the various strategies are far from the optimal coverage. Finally, another interesting avenue for future research is to develop more sophisticated models for compact strategies, such as hybrid models that combine multiple strategies to achieve even better results.

Acknowledgements

The research reported in the paper was partially supported by the PNRR projects “FAIR (PE00000013) - Spoke 9” and “Tech4You

(ECS00000009) - Spoke 6", under the NRRP MUR program funded by the NextGenerationEU.

References

[1] Vincenzo Auletta, Diodato Ferraioli, Valeria Fionda, and Gianluigi Greco, 'Maximizing the spread of an opinion when tertium datur est.', in *Proc. of AAMAS*, volume 19, pp. 1207–1215, (2019).

[2] Vincenzo Auletta, Diodato Ferraioli, and Gianluigi Greco, 'Reasoning about consensus when opinions diffuse through majority dynamics', in *Proc. of IJCAI*, pp. 49–55, (2018).

[3] Vincenzo Auletta, Diodato Ferraioli, and Gianluigi Greco, 'On the complexity of reasoning about opinion diffusion under majority dynamics', *ARTIF INTELL*, **284**, 103288, (2020).

[4] Vincenzo Auletta, Diodato Ferraioli, and Gianluigi Greco, 'On the effectiveness of social proof recommendations in markets with multiple products', in *Proc. of ECAI*, 19–26, (2020).

[5] Vincenzo Auletta, Diodato Ferraioli, and Gianluigi Greco, 'Optimal majority dynamics for the diffusion of an opinion when multiple alternatives are available', *Theor. Comput. Sci.*, **869**, 156–180, (2021).

[6] A. Bhattacharyya., 'On a measure of divergence between two statistical populations defined by their probability distributions', *BULL CALCUTTA MATH S*, **35**, 99–109, (1943).

[7] Sirin Botan, Umberto Grandi, and Laurent Perrussel, 'Multi-issue opinion diffusion under constraints', in *Proc. of AAMAS*, (2019).

[8] Robert Bredereck and Edith Elkind, 'Manipulating opinion diffusion in social networks', in *Proc. of IJCAI*, pp. 894–900, (2017).

[9] Robert Bredereck, Lilian Jacobs, and Leon Kellerhals, 'Maximizing the spread of an opinion in few steps: Opinion diffusion in non-binary networks', in *Proc. of IJCAI*, pp. 1622–1628, (2020).

[10] Matteo Castiglioni, Diodato Ferraioli, and Nicola Gatti, 'Election control in social networks via edge addition or removal', in *Proc. of AAAI*, volume 34, pp. 1878–1885, (2020).

[11] Ning Chen, 'On the approximability of influence in social networks', *SIAM J DISCRETE MATH*, **23**(3), 1400–1415, (2009).

[12] Flavio Chierichetti, Jon Kleinberg, and Sigal Oren, 'On discrete preferences and coordination', *J COMPUT SYST SCI*, **93**, 11–29, (2018).

[13] Dmitry Chistikov, Grzegorz Lisowski, Mike Paterson, and Paolo Turrini, 'Convergence of opinion diffusion is pspace-complete', in *Proc. of AAAI*, volume 34, pp. 7103–7110, (2020).

[14] Gennaro Cordasco, Luisa Gargano, Marco Mecchia, Adele A Rescigno, and Ugo Vaccaro, 'Discovering small target sets in social networks: a fast and effective algorithm', *Algorithmica*, **80**, 1804–1833, (2018).

[15] Morris H DeGroot, 'Reaching a consensus', *J AM STAT ASSOC*, **69**(345), 118–121, (1974).

[16] Pedro Domingos and Matt Richardson, 'Mining the network value of customers', in *Proc. of KDD*, pp. 57–66, (2001).

[17] Piotr Faliszewski, Rica Gonen, Martin Kouteckỳ, and Nimrod Talmon, 'Opinion diffusion and campaigning on society graphs', *J LOGIC COMPUT*, **32**(6), 1162–1194, (2022).

[18] Valeria Fionda and Gianluigi Greco, 'Opinion diffusion in competitive environments: Relating coverage and speed of diffusion', in *COMPLEX NETWORKS*, pp. 425–435, (2019).

[19] M. R. Garey and David S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W. H. Freeman, 1979.

[20] Umberto Grandi, Lawqueen Kanesh, Grzegorz Lisowski, Maadapuzhi Sridharan Ramanujan, and Paolo Turrini, 'Identifying and eliminating majority illusion in social networks', in *Proc. of AAAI*, volume 37, (2023).

[21] M. Granovetter, 'Threshold models of collective behavior', *AM J SOCIOL*, **83**(6), 1420 – 1443, (1978).

[22] D. Kempe, J. M. Kleinberg, and É. Tardos, 'Maximizing the spread of influence through a social network', in *Proc. of SIGKDD*, pp. 137–146, (2003).

[23] David Kempe, Jon M Kleinberg, and Éva Tardos, 'Influential nodes in a diffusion model for social networks.', in *Proc. of ICALP*, volume 5, pp. 1127–1138, (2005).

[24] Yuchen Li, Ju Fan, Yanhao Wang, and Kian-Lee Tan, 'Influence maximization on social graphs: A survey', *IEEE T KNOWL DATA EN*, **30**(10), 1852–1872, (2018).

[25] Vincent Yun Lou, Smriti Bhagat, Laks V.S. Lakshmanan, and Sharan Vaswani, 'Modeling non-progressive phenomena for influence propagation', in *Proc. of COSN*, p. 131–138, (2014).

[26] Ahad N. Zehmakan, 'Majority opinion diffusion in social networks: An adversarial approach', *Proceedings of the AAAI Conference on Artificial Intelligence*, **35**(6), 5611–5619, (May 2021).

[27] Y. Ni, L. Xie, and Z.-Q. Liu, 'Minimizing the expected complete influence time of a social network', *INFORM SCIENCES*, **180**(13), 2514–2527, (2010).

[28] Christos H. Papadimitriou, *Computational Complexity*, Addison Wesley, Reading, MA, USA, 1994.

[29] B. Rozemberczki, R. Davies, R. Sarkar, and C. A. Sutton, 'GEMSEC: graph embedding with self clustering', *CoRR*, (2018).

[30] John P. Scott and Peter J. Carrington, *The SAGE Handbook of Social Network Analysis*, Sage Publications Ltd., 2011.

[31] Paulo Shakarian, Sean Eyre, and Damon Paulo, 'A scalable heuristic for viral marketing under the tipping model', *Soc. Netw. Anal. Min.*, **3**(4), 1225–1248, (2013).

[32] Jaewon Yang and Jure Leskovec, 'Defining and evaluating network communities based on ground-truth.', in *Proc. of ICDM*, pp. 745–754, (2012).

[33] Zhiqiang Zhuang, Kewen Wang, Junhu Wang, Heng Zhang, Zhe Wang, and Zhiguo Gong, 'Lifting majority to unanimity in opinion diffusion', in *Proc. of ECAI*, (2020).