



Stopping fake news: Who should be banned?

Pablo Ignacio Fierens^{1,2} · Leandro Chaves Rêgo^{3,4}

Received: 3 August 2023 / Accepted: 5 March 2024 / Published online: 29 March 2024
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2024

Abstract

Fake news and misinformation spread in online social networks in a manner similar to contagious diseases. One possibility to thwart the contagion cascade is to selectively remove a small number of nodes from the network. Although most of the literature has focused on the selection of those nodes on the basis of their topological position in the network, we pose that attributes of the nodes themselves can be more relevant in certain situations. In order to demonstrate this hypothesis, we introduce a new model of news propagation that accounts for nodes' attributes. In particular, we introduce three important characteristics of a node: the influence capacity, the resistance to be influenced and the resistance to become an information spreader. Besides offering an intuitive justification for the model and these new parameters, we relate them to other proposals in the literature. Under the new model and using numerical simulations on both synthetic and real life networks, we show that nodes' attributes can be more important than their graph structural properties in choosing an adequate set of vertices to be removed with the purpose of mitigating fake news propagation. Furthermore, our results suggest that removal of nodes with high influence power is more effective in denser networks and when the influence of a few nodes is much larger than that of the general population.

Keywords Complex networks · Misinformation · Information diffusion · Online social networks

1 Introduction

We are all familiar with the existence of people who believe that the Earth is flat [1], or those that, without solid scientific evidence, firmly believe that we have been visited by extrater-

restrial beings [2]. Furthermore, we know that in an era of social networks like Facebook, Instagram, Twitter, Twitch or YouTube, this type of ideas propagates more easily [3, 4]. Probably, most of us do not mind about these conspiracy theories. After all, it is a matter of *beliefs*: there is no amount of scientific evidence that could help to change the mind of a believer.

However, conspiracy theories may sometimes become dangerous. As an example, many governments and non-governmental organizations around the world strongly fought against anti-vaccination movements in the midst of the worldwide COVID-19 pandemic [5–10]. Although education and publicity campaigns based on verified information may help to oppose the spread of misinformation on the much needed vaccines, as in the case of the flat-earthers, believers are difficult to convince even in the face of abundant scientific data.

Some tech companies, such as Facebook [11, 12] and Google [13], that run social networking applications have announced several measures to combat false news. There is also a large body of research in this area. However, most of the scientific literature and company-based initiatives focus on either the automatic or man-in-the-loop detection of false news in social media [14–17]. In this work, to the contrary,

Pablo Ignacio Fierens and Leandro Chaves Rêgo have contributed equally to this work.

✉ Pablo Ignacio Fierens
pfierens@itba.edu.ar

Leandro Chaves Rêgo
leandro@dema.ufc.br

¹ Instituto Tecnológico de Buenos Aires (ITBA), Iguazú 341, C1437ETC Ciudad Autónoma de Buenos Aires, CABA, Argentina

² Department, National Scientific and Technical Research Council (CONICET), Godoy Cruz 2290, C1425FQB Ciudad Autónoma de Buenos Aires, CABA, Argentina

³ Departamento de Estatística e Matemática Aplicada, Universidade Federal do Ceará, Bloco 910, Campus do Pici, Fortaleza, Ceará 60440-554, Brazil

⁴ Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Pernambuco, Av. Prof. Moraes Rego, 1235, Cidade Universitária, Recife 50.740-550, Brazil

we focus on a different aspect of misinformation: the network that facilitates its spread.

At the beginning of the year 2021, we witnessed a very important event: a large social network, Twitter, banned the account of the United States president [18, 19]. In spite of the political content of this measure, it signals to an alternative in the combat against misinformation and false news: the ban of misinformation super-spreaders. ‘Super-spreader’ is a term that has been introduced, at least for the general public, in the last few years in relation to the COVID-19 pandemic. A super-spreader is a person with a high viral load that gets in contact with large number of people, for example, in parties or other social gatherings (weddings or religious events). By analogy, the concept has been transferred to the realm of information. Indeed, recently some presidents of large countries have been called super-spreader of disinformation [20, 21]. The concept has caught up and it is becoming part of common language in the media [22, 23].

As a matter of fact, the idea of super-spreader is not new. It is related to the well-known concept of an opinion leader or ‘influential’ [24]. Opinion leaders are defined as ‘the individuals who are likely to influence other persons in their immediate environment’ [24]. As it is clearly explained in Watts and Dodds [24], an opinion leader is not necessarily a leader, such as the president of a country. A simple example that comes to mind is that of a company that wants to sell plastic containers. Opinion leaders in this case can be certain well-connected housewives, and not a high-profile person. In this sense, the concept of an influential is different from the nowadays common concept of ‘influencer’ [25–29], that is, someone ‘who built a large network of followers, and are regarded as trusted tastemakers in one or several niches’ [28].

Super-spreaders, influentials or opinion leaders, and influencers are all related, but slightly different concepts. A common factor in all of them is that of the network of connections. Super-spreaders ‘spread’ misinformation or false news because they have a large number of connections in propitious environments (as those who spread the COVID-19 infections in large gatherings in closed rooms). Influentials can be characterized by their special place in a network of contacts [24, 30–33]. Influencers are defined by their large network of ‘followers’ [28]. All in all, the network topology is in the core of the definition of these types of characters relevant in opinion formation dynamics.

A careful reading of the previous lines reveals a detail which is of the utmost relevance. Indeed, although the topology of connections in a social network is important in the definition of either a super-spreader or an influencer, that is not the only aspect that characterizes them. For example, a super-spreader not only needs a large number of connections, but also a propitious ‘environment.’ Besides being followed by a large number of people, an influencer requires to be ‘regarded as a trusted tastemaker’ [28]. In summary, there is

an attribute of the super-spreader or the influencer that goes beyond their place in the social network.

The main goal of this work is to investigate whether there may exist attributes of individuals that can be more relevant than their place in the network topology when evaluating the importance of their influence. If there are such attributes, they can help to detect and combat super-spreaders, which may be more efficient than fighting the fake news they divulge. A related idea was treated in Andrade et al. [34], where it was studied how to efficiently choose what elements of a criminal network should be targeted to cause the largest damage. It was found that not only the connections in the network were important, but also attributes of the criminals. In a similar manner, we propose to identify who are the news-spreaders (influencers) that are more convenient to combat in order to cause the greatest damage to a misinformation network.

A misinformation network is different than a criminal network. There is a large literature on opinion formation and its relation to a network of contacts [24, 35–46]. Oftentimes, fake news are associated with the idea of an spreading virus. Indeed, it is common to speak of a false news becoming ‘viral.’ A classical model for the propagation of a disease is known as the susceptible–infected–recovered (SIR) model [24, 47, 48]. The basic idea of this model is to classify people in three categories: ‘susceptible,’ they can become infected; ‘infected,’ as they are sick and can infect others; and ‘recovered,’ as they were but are no longer sick. Susceptible individuals can be infected if they come into contact with infected individuals. There is a probability of infection which may depend on several factors, but mainly on the type of contact (for example, how long a susceptible individual was in contact with an infected person). Sick (infected) people recover at a constant mean rate.

It is most interesting to study the behavior of the SIR model in the context of a social network [24]. There have been several proposals on how the probability of contagion is related to the type of connection between a susceptible individual and an infected person. During the COVID-19 pandemic, we have grown used to terms such as ‘close contact.’ The closeness of the contact is an attribute of the link between the two individuals. However, we argue that the contagion probability also depends on characteristics of the infected and susceptible individuals beyond the properties of the connection. As in the example of presidents of large countries as super-spreader of (dis)information, there are attributes of the individuals themselves which are relevant. Indeed, already Katz [49] recognized three dimensions related to an influential or opinion leader (we follow Nisbet and Kotcher [50]): (i) ‘who one is’: certain personality traits; ‘what one knows’: expertise about a particular subject; and ‘whom one knows’: number of contacts. While the last dimension can be related to topological properties of the network of contacts, the first two are connected to personal attributes of the particular node. In relation

to dimension (i), we can mention, for example, personality strength [51] and communication abilities [52]. Authoritative individuals are not always characterized by their knowledge (dimension (ii)), but also by their social standing in a given community [53, 54]. Another important attribute that has recently been studied is the personal motivation [55, 56]. For more information on individual attributes that may be relevant, the interested reader is referred to the works of Valente and Pumpuang [57] and Bamakan et al. [58] and references therein. In this work, we study news propagation in a Twitter network, where the number of mentions is used as a node attribute.

Our aim is to find a way of identifying the misinformation super-spreaders (influencers) which are more efficient to combat in order to produce the greatest damage to a misinformation propagation. In order to evaluate this, we propose a novel approach to the analysis of the propagation of fake news which, on the basis of a variation of the SIR model, takes into account both the topology of the social network and the attributes of the nodes.

The remaining of the paper is organized as follows. Section 2 presents a review of some of the most common (network-structure related) properties of nodes considered in the literature. In Sect. 3, we put forth a model of information diffusion that accounts for the influence power of the agents. We also discuss the relation of our model with related literature on complex contagions. Section 4 studies the behavior of the model and the impact of each of its parameters by means of simulations on both synthetic and real-world networks. Finally, we close the paper with some conclusions in Sect. 5.

2 Topological properties

Relevance of nodes is usually appraised through different centrality metrics which only consider their network-structure related characteristics. In this section, we review some of the most common centrality metrics. However, as we shall show in Sect. 4, they are not always appropriate in the context our proposed information diffusion model.

Let us consider a social network $G = (V, E)$ where V is the set of vertices or nodes, $E \subseteq V \times V$ is the set edges between the nodes. For the sake of clarity, in what follows we shall assume that all nodes are uniquely identified by a natural number in $\{1, 2, \dots, n\}$, where n is the number of vertices in the network. Edges may be either directed or undirected, although for our application directed edges are usually more appropriate. Indeed, an online network such as Instagram has many important influencers that are ‘followed’ by many users which are not necessarily followed-back. In the context of a graph, following the flow of information, we call predecessors of a node to those nodes it follows and successors to its followers. We must note that since $E \subseteq V \times V$, only one-

to-one interactions are considered in our modeling approach. However, higher-order interactions can be of great interest in the propagation of information or diseases [59, 60].

The simplest node centrality measure is its degree, i.e., the number of links to other nodes. In a directed network, we can distinguish between in (d^{in}) and out degree (d^{out}) depending on whether we account for incoming or outgoing links, respectively. In the context of information propagation, d^{out} appears to be more adequate as it refers to the news-spreading capacity of the node.

A path from node i to j in a network consists of a sequence of edges $(i_{k-1}, i_k) \in E$, for $k = 1, 2, \dots, p$, where $i_0 = i$ and $i_p = j$. A geodesic path from node i to node j is a path between them with the smallest number of edges in the sequence and this smallest number is called geodesic distance between i and j and denoted by d_{ij} . For a given node i , let V_i be the set of nodes $j \neq i$ in the network for which there exists a path from i to j . Let n_i be the number of elements in V_i . Closeness centrality [61] measures the reciprocal of the mean distance from a node to other nodes in the network that it can reach by a path:

$$c_{cl}(i) = \frac{n_i}{\sum_{j \in V_i} d_{ij}} \tag{1}$$

We must remark that distance can be measured in other ways, e.g., incorporating edge-related properties. However, in the context of the model presented in the following section, we consider the geodesic distance. Intuitively, we may expect that, the closest a node is to all other nodes, the fastest a news will reach all nodes that can be reached from that node.

While the degree only considers how many nodes can be reached in a single hop, the rapid spread of a news also depends on the number of nodes connected by the first layer of neighbors. In this sense, eigenvector centrality [62] appears to be an adequate measure as the relevance of the node is proportional to the sum of the centralities of its neighbors. Let \mathbf{A} be the adjacency matrix of the network G , i.e., a matrix such that $\mathbf{A}_{ij} = 1$ if there is a link from node i to node j , and $\mathbf{A}_{ij} = 0$ otherwise. The eigenvector centrality is defined by the vector that satisfies

$$\lambda_1 \vec{c}_{eg}^T = \vec{c}_{eg}^T \mathbf{A}, \tag{2}$$

where λ_1 is the largest left eigenvalue of matrix \mathbf{A} and the eigenvector centrality of node i , $c_{eg}(i)$, is given by the i th element of vector \vec{c}_{eg} . Observe that we are using left eigenvectors as the centrality of a node depends on the centralities of its successors, as it is more relevant in the new-spreading context.

A measure related to eigenvector centrality is PageRank [63], originally devised to evaluate the importance of webpages. In terms of the adjacency matrix, we can write

the vector of PageRank centralities as [61]

$$\vec{c}_{pr}^T = (1 - \alpha)\vec{1}^T (\mathbf{I} - \alpha\mathbf{AD})^{-1}, \tag{3}$$

where \mathbf{I} is the $n \times n$ identity matrix, n is the number of nodes in the network, and \mathbf{D} is a diagonal matrix such that $\mathbf{D}_{ii} = \max\{1, d^{\text{out}}(i)\}$. The parameter $\alpha \in (0, 1)$ is conventionally assumed equal to 0.85.

Betweenness [64, 65] is a centrality measure which quantifies the importance of a node for the flow of information in a network. In particular, the betweenness centrality of node i is defined by

$$c_{be}(i) = \sum_{j,k \neq i} \frac{\text{\#of shortest paths from } j \text{ to } k \text{ that go through } i}{\text{\#of shortest paths from } j \text{ to } k}. \tag{4}$$

Although there are many other centrality measures, we shall focus only on these ones, as they are the most common. For metrics with application to information diffusion, we refer the interested reader to, e.g., Kitsak et al. [30], Pei and Makse [31], Pei et al. [32], and Taha [33].

3 Proposed model

We introduce a new information diffusion model which is based on the well-known susceptible–infected–recovered (SIR) contagion model [24, 47, 48]. In particular, we consider that each node can be in any of three states:

- Susceptible (S): it may or may not have received the information, but the vertex still does not believe in it.
- Infected (I): it is aware and believes the information, but it still does not propagate it to its connections.
- Spreader (P): it actively propagates the news to all of its successors in the network.

The distinction between an infected and a spreader node has already been proposed in the literature of information diffusion (see, e.g., Xiong et al. [66]). Note that we are not considering the possibility of a recovery, i.e., in this model an infected vertex cannot return to the susceptible state. This implies that we are not allowing an opinion change; that is, once a node is infected and believes some information, it cannot disbelieve it in the future. Even though this choice might seem unrealistic, it simplifies the analysis. Thus, we leave modeling opinion changes as a subject of future work.

Let us consider an online social network such as Instagram or Twitter. A spreader posts a comment or shares another post commenting a news. When will its followers become aware of the post? The answer of this question is complex,

as it depends on many factors such as details of the online social network (e.g., the order in which posts are presented), habits of the successors, etc. For this reason, in this paper we assume, as a first approximation, a constant mean rate of reads of the news. In particular, the time between the moment a spreader posts a message and a follower reads it is modeled by an exponential random variable with constant parameter λ .

A person may have the potential to be very influential. For example, he/she may be very charismatic, a great speaker or writer, or an excellent producer of memes. We characterize the spreading ability of node i by the parameter $\beta_i^{\text{max}} \in \mathbb{R}^{\geq 0}$. However, that particular person may not use all of its potential, e.g., it may depend on his/her current involvement with the subject. The influence capacity of node i at time t is modeled by $\beta_i(t) \in \mathbb{R}^{\geq 0}$, which is always smaller than or equal to β_i^{max} . The actual current influence capacity of node i is determined by the influence power of those who have contacted it:

$$\beta_i(t) = \beta_i^{\text{max}} \times \Psi_i \left(\sum_{t_{ji} \leq t} \beta_j(t_{ji}) \right), \tag{5}$$

where $\Psi_i : \mathbb{R} \rightarrow [0, 1]$ is a non-decreasing function and t_{ji} is the time when node i read the information published by its predecessor node j . Intuitively, the involvement of vertex i with the news and, hence, its spreading power depends on the strength with which node i received the information from its predecessors. Although more complex alternatives are possible, for the sake of simplicity we assume that

$$\Psi_i(x) = \begin{cases} 0 & x < \psi_i, \\ \frac{x - \psi_i}{\Delta_i} & \psi_i \leq x \leq \psi_i + \Delta_i, \\ 1 & x > \psi_i + \Delta_i, \end{cases} \tag{6}$$

where $\psi_i, \Delta_i \in \mathbb{R}^{\geq 0}$ are two parameters. While the value of ψ_i quantifies the ‘resistance’ of node i to become a spreader, Δ_i specifies how quickly it achieves its maximum influence power. Melnik et al. [67] also considered nodes with different influence capacities. However, their influence was discretized in a finite set of levels.

Susceptible nodes become infected if they have been exposed to sufficiently strong influences. In particular, node i goes from state S to state I after being informed of the news by node k if

$$\sum_{t_{ji} \leq t_{ki}} \beta_j(t_{ji}) \geq \phi_i, \tag{7}$$

where ϕ_i quantifies the resistance of node i to believe any piece of information. Since we assume that a node must be infected, i.e., it must believe in the information, before

spreading it, we require $\psi_i \geq \phi_i$. It must be noted, though, that there may be situations in which it might be convenient to assume that a spreader does not necessarily believe the news. A related approach was proposed by Huang et al. [68] that distinguished between infected and persuader nodes, with different thresholds to enable the transition from the susceptible state. In their setting, all infected nodes are propagators, but persuader agents have a larger influence power as they ‘immediately’ infect (persuade) all their neighbors.

It is important to observe that a node might need to be exposed to the same news from different sources in order to believe it or to start spreading it. However, this number depends on the influence capacity of its predecessors ($\beta_j(t_{ji})$) and the resistance of the node to believe (ϕ_i) and propagate the news (ψ_i). Contagion paradigms where exposures from more than a single neighbor are needed in order to become infected have been dubbed in the literature as *complex contagion models*. The seminal work by Centola and Macy [69] considered the case where a node became infected only if the fraction of infected neighbors was greater than a threshold. While Centola and Macy [69] focused on the structural properties of networks that facilitate or hinder contagion cascades, it also studied the impact of network heterogeneity and even the existence of some higher-status nodes (in the context of our model, nodes with high β_i^{\max}). In this sense, mathematical details aside, our proposal introduces a new concept, that of the involvement of a node in the propagation of a news as dependent on the status of those from whom it received it. This approach can be related to the observation that the propagation of a piece of information depends on the strength of the link on which that information was received [70]. However, let us observe that, in our model, the influence power is a property of the node and not of the link between nodes.

There is another significant difference between this paper and most of the literature on complex contagions. Indeed, the vast majority of the papers are focused on the modeling side [68–78] or, in social and behavioral change applications, concerned in the best seeding strategies to diffuse a new cultural paradigm [79–85]. On the contrary, we are interested in how to stop or hinder the diffusion of a news. Profound analysis of how information propagates can shed light on how to combat the spread of news [86]. However, we turn to a more straightforward approach by looking for the nodes that should be combated in order to most efficiently thwart diffusion. In the context of complex contagions, this approach has already been explored by Centola [87] that studied the robustness of diffusion in scale-free and exponential networks. However, Centola only analyzed the size of the infection cascade and the influence of removing either the highest degree or randomly chosen nodes. Kuhlman et al. [88] extended the work of Centola by analyzing other criteria for node selection. They showed that, under their modeling framework, selecting

a minimal set of nodes to block infection is NP-hard and proposed efficient heuristics. Their work was further extended for more than one ‘illness’ in Carscadden et al. [89]. We must remark that all these works do not consider the influence potential of the nodes as we do in this paper. In general, the problem of finding an optimal set of nodes to be removed is dubbed as the *critical node detection problem* and there is a vast literature on the subject. The interested reader is referred to the review by Lalou, Tahroui and Khedouci [90]. It is interesting to note, however, that most papers consider that nodes can actually be removed from the network. Cavallo et al. [91] argue that this may not be always the case, as some nodes may resist to be removed. While do not pursue this idea further in this work, it deserves to be studied in the future.

A simple alternative to stop news diffusion is to attack the first infected nodes, but in practical settings those nodes might be discovered only when the infection has started to propagate and, thus, it would be inefficient as an immediate reaction. Moreover, in an online social network, new infections may come from external sources [70, 92] out of the control of the social network itself. Therefore, if it does not make sense to look for the first news-spreaders, what other nodes should be attacked or even banned from network? This is the fundamental question we try answer in this work.

4 Numerical simulations

Since the time between the moment a spreader posts a message and a follower reads it is modeled by an exponential random variable with constant parameter λ , our simulations are based on Gillespie’s algorithm [93]. In particular, the time to the next infection, δT , is simulated as

$$\delta T = -\frac{\ln(r_1)}{\sum_{j=1}^n \sum_{i \in \mathcal{P}_j} \lambda}, \quad (8)$$

where r_1 is a random number in $(0, 1)$, \mathcal{P}_j is the set of predecessors of j that are spreaders and have yet not infected j . We determine which node k is infected by generating a new random number r_2 with a uniform distribution in $(0, 1)$ and looking for the minimum k such that

$$\sum_{j=1}^k \sum_{i \in \mathcal{P}_j} \lambda \geq r_2 \times \sum_{j=1}^n \sum_{i \in \mathcal{P}_j} \lambda. \quad (9)$$

Finally, we assume an ordering of \mathcal{P}_k and choose the infecting node as the first one such the sum on the left is greater or equal than that on the right.

In order to understand the workings of the proposed model, we begin by studying its behavior in synthetic net-

Table 1 Simulation parameters—Barabási–Albert network

Parameter	Notation	Value
Number of nodes	n	1000
Number of edges of each new node	m	5
Influence capacity	β_i^{\max}	$U(0.9, 1.1)$
Special influence capacity	β_{gr}	1000
Resistance to becoming infected	ϕ_i	$U(0.9, 1.1)$
Resistance to becoming a spreader	ψ_i	$U(\phi_i, 1.1)$
Resistance to attain the maximum influence power	Δ_i	4
Number of nodes with special influence capacity	n_{gr}	10
Number of nodes removed	n_{rem}	10

$U(a, b)$ corresponds to a uniform distribution between a and b

works of the type of Barabási–Albert [94]. In particular, we start with a network with $n = 1000$ nodes where each new added node attaches to other $m = 5$ vertices according to the preferential attachment mechanism. Since we shall focus on networks where most of the nodes are similar and only a few of them stand out as great potential influencers, we assign attributes randomly to each vertex according to the following rules: $\beta_i^{\max} \sim U(0.9, 1.1)$, $\phi_i \sim U(0.9, 1.1)$, $\psi_i \sim U(\phi_i, 1.1)$, where $U(a, b)$ corresponds to the uniform distribution between a and b . The choice of the distribution is not significant, and it only introduces a small heterogeneity among nodes. For the sake of reference, we summarize all simulation parameters in Table 1.

Our interest is to find which nodes are more convenient to remove from the network to either stop or alleviate the propagation of the news. We shall evaluate the convenience of removing nodes according to some topological metric such as the out degree, the closeness centrality, the eigenvector centrality, the PageRank centrality, and the betweenness centrality. We also study the benefit of removing nodes with largest influence potential β_i^{\max} and, for the sake of comparison, randomly chosen vertices. After removing 10 nodes based on one of these criteria, we choose other 10 nodes at random to become initial spreaders with its maximum influence power, i.e., $\beta_i(0) = \beta_i^{\max}$. While the choice of the number of removed nodes is, in principle, arbitrary, it must be observed that it agrees with n_{gr} , the number of nodes with $\beta_i^{\max} = \beta_{\text{gr}}$. That is, if we remove some of the more influential nodes, we remove all of them. By the end of this section we consider the influence of this decision by varying n_{gr} . We must also mention that the set of firstly infected nodes may or may not have intersection with the set of more influential nodes. This choice is in agreement with previous studies [95].

Figure 1 presents boxplots showing the first three quartiles of the number of infected and spreaders. The results correspond to 10,000 simulations with different networks. We used $\lambda = 1$ and the panel on the left shows the number of I and P after 2 time units, while the right panel shows the

numbers after an infinite time period. We do not present the outliers in the boxplot graph for the sake of clarity. In principle, there does not appear to have any clear advantage using any of the proposed metrics and, in general, random removal of nodes is worst. Thus, in the remaining of the paper and for the sake of clarity, we shall focus our attention on the removal of nodes with the highest out degree d_i^{out} and with the highest influence potential β_i^{\max} .

The temporal evolution of the mean and the standard deviation of the number of infected and spreader nodes is shown in Fig. 2. As it can be observed, removing the vertices with highest β_i^{\max} not only is the best option on the average, but it also exhibits the lowest standard deviation. Nonetheless, the spread of values is such that it might be possible that, in a number of situations, removing the nodes with largest degree might result advantageous. Since the large standard deviation appears to be inherent to the Poissonian contagion processes, we shall focus on the average behavior in the remaining of this work.

It is interesting to observe the influence of the network density in the temporal evolution determined by our model. Figures 3 and 4 show results for Barabási–Albert networks where each new added node links to $m = 2$ and $m = 10$ existing nodes, respectively. As it can be observed, removing nodes with higher degree is more effective than removing vertices with higher influence potential when the density of the network is low. The intuition behind this result comes from the fact that super-spreaders need their neighbors to be well connected so that their influence spreads through the network. Therefore, networks with higher density are more susceptible to damage caused by nodes with high influence power.

In order to understand the impact of each parameter in the model, we vary them while showing the results after 2 time units for the case of the Barabási–Albert graph with $m = 5$ new links with each new node. The left panel of Fig. 5 shows that, as expected, the removal of the nodes with higher β_i^{\max} becomes more effective as the value of β_{gr} increases. We can

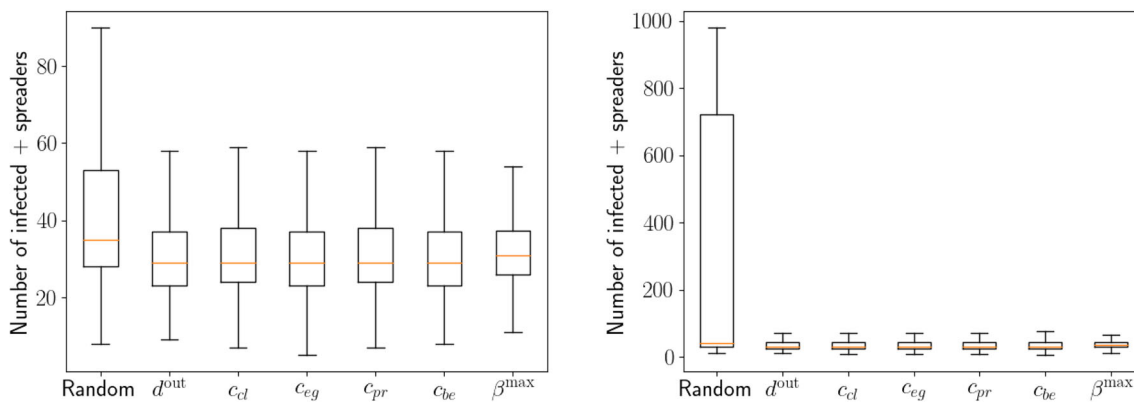


Fig. 1 Boxplots for the number of infected and spreader nodes after 2 time units (left) and after an infinite amount of time (right). Outliers are not presented for the sake of clarity. Each boxplot corresponds to the result after the removal of randomly chosen nodes (Random), vertices

with the highest out degree (d^{out}), closeness (c_{cl}), eigenvector centrality (c_{eg}), PageRank (c_{pr}), betweenness (c_{be}), and influence potential (β^{max})

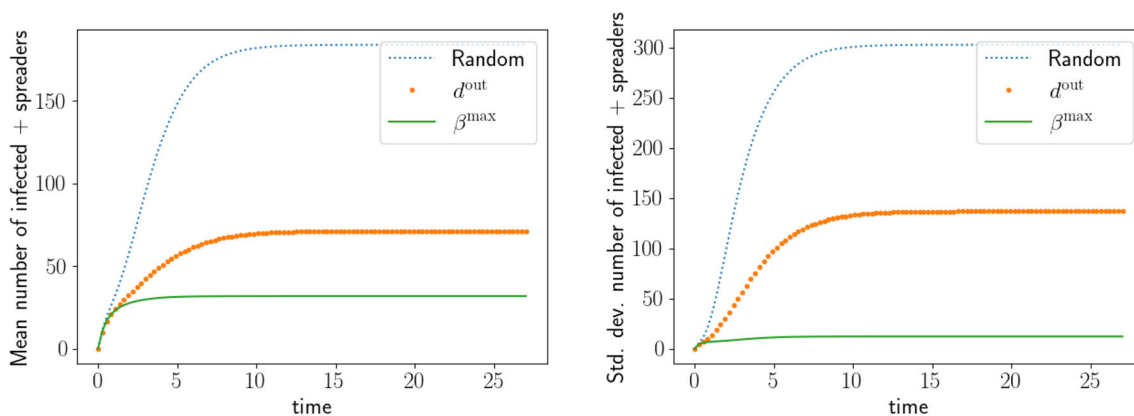


Fig. 2 Temporal evolution of the mean (left) and standard deviation of the number of infected and spreader nodes in Barabási–Albert networks with $m = 5$ connections made by each new added node

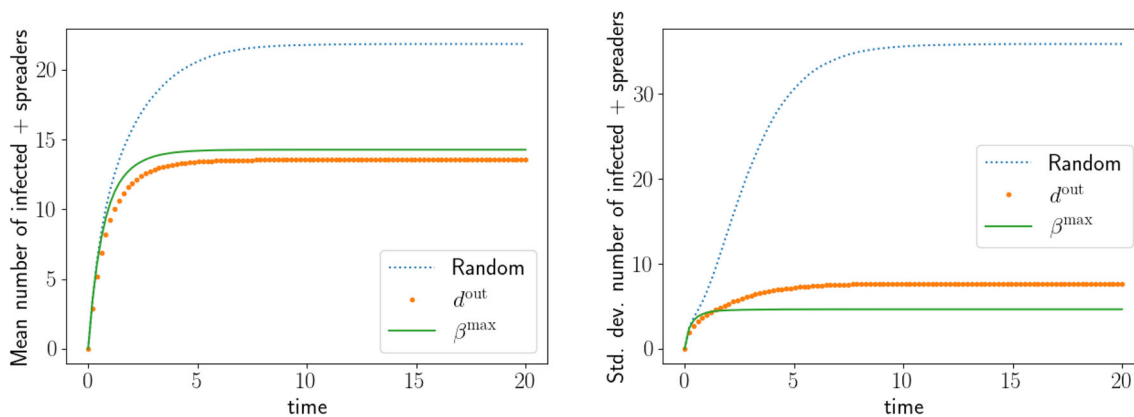


Fig. 3 Temporal evolution of the mean (left) and standard deviation of the number of infected and spreader nodes in Barabási–Albert networks with $m = 2$ connections made by each new added node

arrive at a similar conclusion in relation to changes to Δ_i , as shown in the right panel of Fig. 5.

The left panel of Fig. 6 presents results when the mean value of ϕ_i is changed. These results correspond to assigning the attributes $\phi_i \sim U(\bar{\phi}-0.1, \bar{\phi}+0.1)$, $\psi_i \sim U(\phi_i, \bar{\phi}+0.1)$.

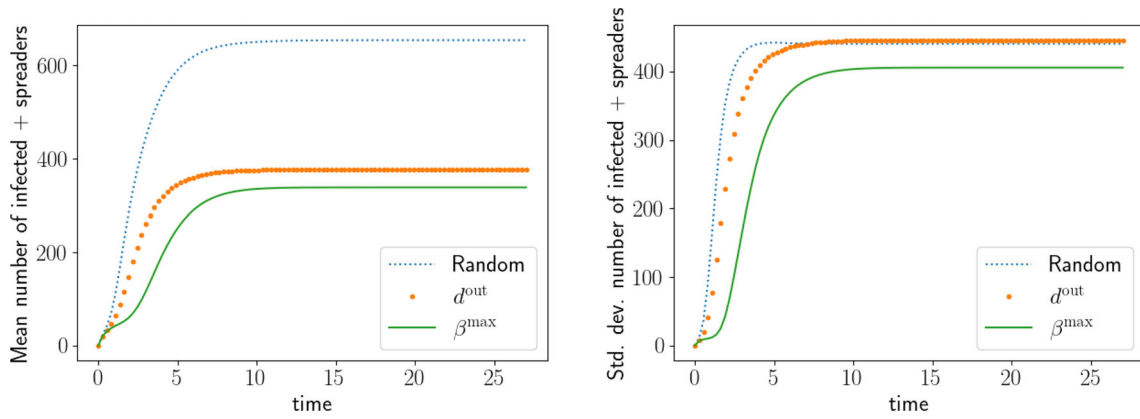


Fig. 4 Temporal evolution of the mean (left) and standard deviation of the number of infected and spreader nodes in Barabási–Albert networks with $m = 10$ connections made by each new added node

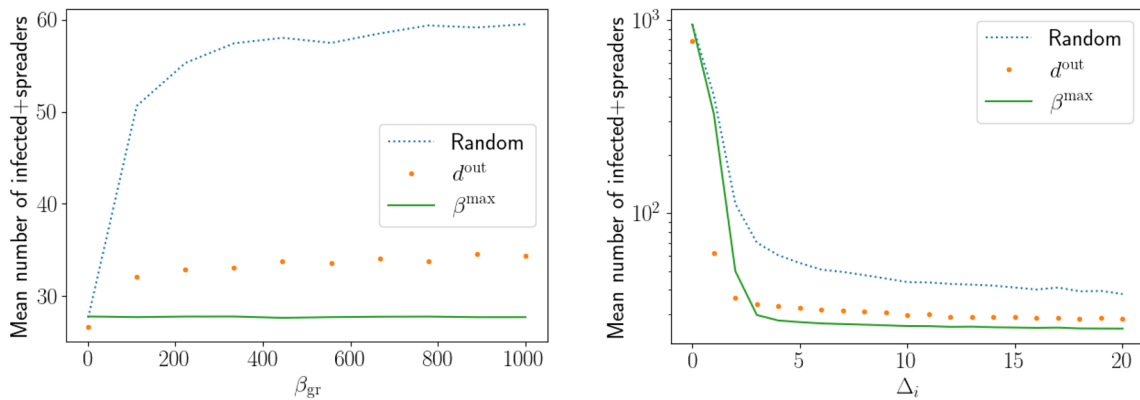


Fig. 5 Influence of β_{gr} (left) and Δ_i (right). A logarithmic scale is used in the second figure in order to better appreciate the differences between the three curves

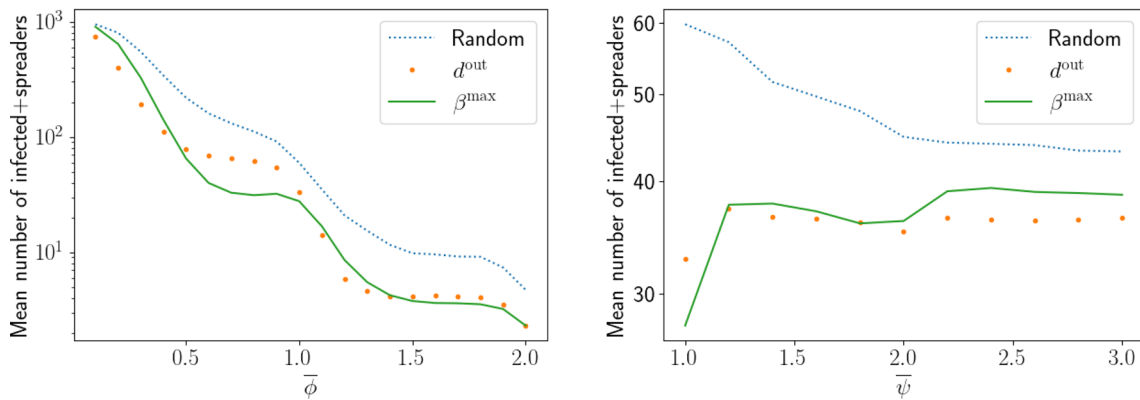


Fig. 6 Influence of $\bar{\phi}$ (left) and $\bar{\psi}$ (right). A logarithmic scale is used in the first figure in order to better appreciate the differences between the three curves

It is immediate to observe that as the resistances of the nodes to become infected and spreader increase, the number of infected vertices decreases. For small values of ϕ_i ($\bar{\phi} < 0.5$), nodes are easily infected and, thus, it is convenient to remove higher degree nodes as there is no need for a large influence power for the spread of misinformation. How-

ever, when resistance to infection increases ($\bar{\phi} \in (0.5, 1.0)$), influence power of spreading nodes becomes more relevant and it is more effective to remove nodes higher β^{\max} . Finally, for large resistance values ($\bar{\phi} > 1$), there is not a significant difference between the behavior β when the nodes of higher degree or the nodes with largest β^{\max} are removed. Since the

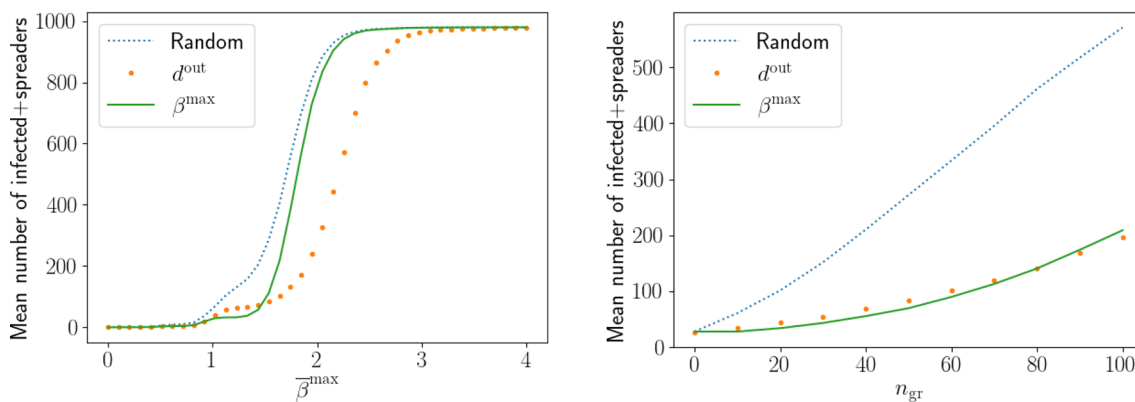


Fig. 7 Influence of $\bar{\beta}_i^{\max}$ (left) and of the number of nodes with $\beta_i^{\max} = \beta_{gr}$ (right)

initially infected nodes are a small portion of the population and the resistance to infection is high, it may be the case that the infection does not reach nodes with high degree or large β^{\max} by the time considered in Fig. 6 (2 time units).

Results where only the value of ψ_i is changed are shown in the right panel of Fig. 6. In particular, nodes received the value $\psi_i \sim U(\max(\bar{\psi} - 0.1, \phi_i), \bar{\psi} + 0.1)$. Observe that, for higher values of $\bar{\psi}$, it becomes more effective to remove nodes with higher degree. As a matter of fact, once the nodes with higher β^{\max} have spread the information as much as they can (if they are present), most of the contagions are due to ‘regular’ nodes. If the resistance to become a spreader increases, there is a need for a larger number of contagions to transform an infected node into a spreader. A larger number of contagions is possible if the node has a high in degree. Since the Barabási–Albert network is, in effect, undirected, removing the nodes with high degree hinders this process.

We vary the influence potential of regular nodes, by assigning $\beta_i^{\max} \sim U(\bar{\beta}^{\max} - 0.1, \bar{\beta}^{\max} + 0.1)$ and changing the mean value $\bar{\beta}^{\max}$. As it can be observed in the left panel of Fig. 7, the removal of the nodes with higher β_i^{\max} becomes less effective as $\bar{\beta}^{\max}$ increases. Intuitively, as the regular nodes become potentially better spreaders, the importance of nodes with large influence power diminishes in comparison.

We also evaluate the influence of n_{gr} , the number of higher-status nodes with large $\beta_i^{\max} = \beta_{gr}$. Keeping constant in 10 the number of removed nodes and the number of initially infected nodes, as the number of nodes with $\beta_i^{\max} = \beta_{gr}$ increases, it becomes less convenient to remove them, in comparison with the removal of nodes with high degree. As it can be observed in the right panel of Fig. 7, when $n_{gr} > 80$, the removal of high-degree nodes is more efficient. This suggests that the removal of high β_i^{\max} is more important when a large portion of them is removed from the network. If this is not possible, then it is better to focus on the removal of high-degree nodes.

In previous simulations, the value of the influence power β_i^{\max} was uncorrelated with the degree. Although this is a convenient starting point of analysis, in real-world scenarios a relation between both quantities is to be expected. In order to understand the consequences of a correlation between them, we set up networks as those used in Fig. 2. However, the nodes with $\beta_i^{\max} = \beta_{gr}$ were not uniformly chosen at random from the set of all vertices, but with a probability proportional to $(d_i^{\text{out}})^\alpha$. The value of $\alpha \in \mathbb{R}$ changes the correlation between the out degree and the influence power, but it does not alter their marginal distributions. Figure 8 shows the mean (left panel) and the standard deviation (right panel) of the number of infected nodes at time $t = 5$, for 10,000 simulations and $\alpha \in [-10, +10]$. As it can be readily observed, there is a range of values of α , corresponding to both negative and positive correlations, for which the deletion of the nodes with higher influence capacity is the more convenient option. For $\alpha < -2$, nodes with large β_i^{\max} are poorly connected and, thus, it is more convenient to isolate vertices with higher degree. On the contrary, for $\alpha > 3$, nodes with higher degree are also (with high probability) nodes with high influence power. Hence, isolating either with respect d_i^{out} or β_i^{\max} yield almost the same results.

4.1 Twitter network

The use of synthetic networks has the advantage of allowing control their characteristics; however, their structures may be somewhat different from that of networks found in the wild. Moreover, the Barabási–Albert graph is intrinsically undirected and we are more interested in the directed propagation of information. For these reasons, in this section we apply our model to real network data. In particular, we use data from Twitter that was first presented in De Domenico et al. [96]. The dataset was built by following the messages posted in Twitter about the discovery of the Higgs boson between 1 and 7 July 2012, and it contains the network of

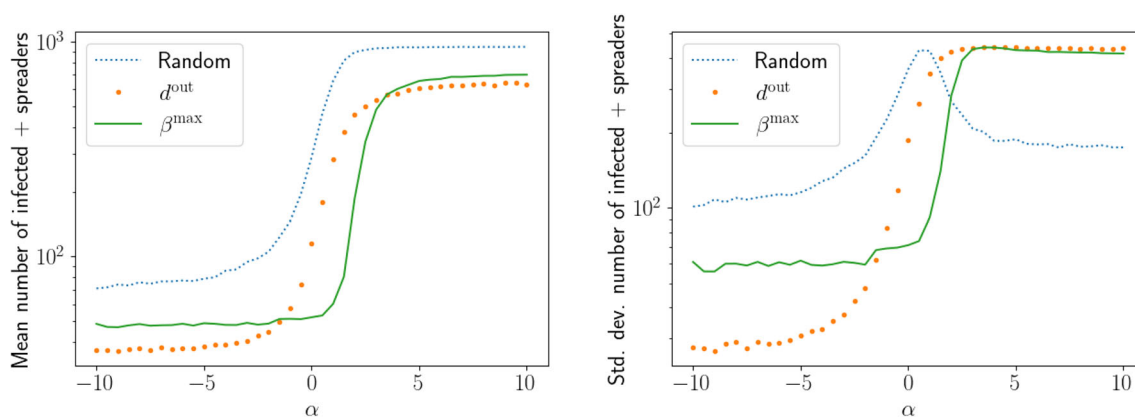


Fig. 8 Mean (left) and standard deviation (right) of the number of infected and spreader nodes in Barabási–Albert networks with $m = 5$ connections made by each new added node. Nodes with $\beta_i^{\max} = \beta^{\text{gr}}$ were chosen with a probability proportional to $(d_i^{\text{out}})^\alpha$

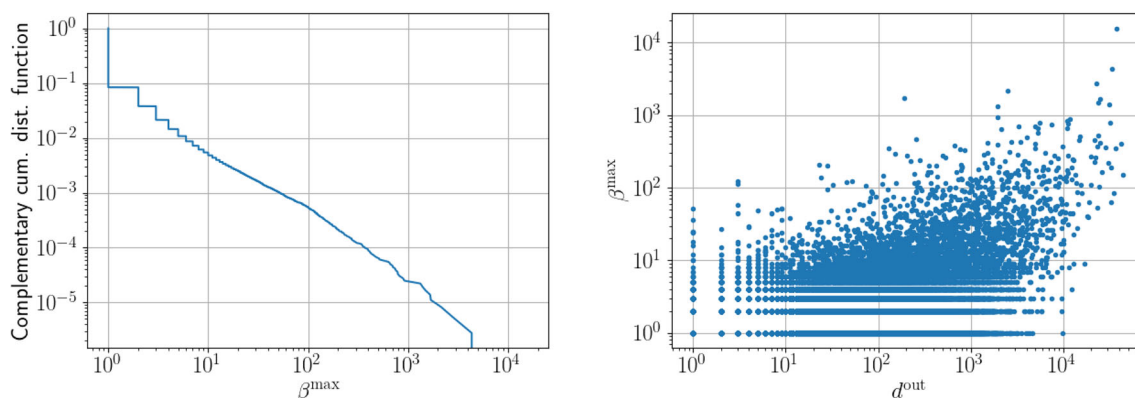


Fig. 9 Complementary cumulative distribution function of the spreading capacity (left) and capacity as a function of out degree (right) in the Twitter network

Twitter followers as well as statistics on retweets, replies and mentions.

From the social graph, we kept only the largest strongly connected component consisting of 360,210 vertices and 14,102,583 directed edges. In order to represent the flow of information, edges were directed as going from accounts to each of their followers. One very interesting characteristic of the dataset is that there are several non-topological properties that may be considered as measures of influence power; for example, the number of times an account was mentioned or the number of retweets of the account’s posts. Thus, let us assume the number of mentions n_i^{men} that node i received as corresponding to its spreading capacity β_i^{\max} . However, given the short period of time (7 days) that was surveyed, not every node was mentioned. In order to allow for each node to have some spreading capacity, we set $\beta_i^{\max} = n_i^{\text{men}} + 1$. The left panel of Fig. 9 shows the complementary cumulative distribution function of the spreading capacity, showing that a large fraction of the nodes has a small value of β^{\max} .

It is interesting to analyze the relation between the spreading capacity of each node in the resulting network with some

its topological properties. The right panel of Fig. 9 presents a scatter plot of the spreading capacity as a function of the out degree. The correlation coefficient between both quantities is ~ 0.37 , where its positivity is related to the observed tendency of higher β^{\max} to correspond to larger d^{out} . However, the correlation coefficient is not close to unity and the popularity of a node, measured here by the number of mentions, is not perfectly correlated with a topological property such as its degree. This fact is on the basis of our proposal, i.e., that there may exist some non-topological properties of vertices which are important in the dissemination of news.

Since some nodes reach values of $\beta^{\max} > 10^4$, we set a large value of Δ_i , namely, $\Delta_i = 100$. The remaining parameters are randomly assigned: $\phi_i \sim U(0.9, 1.1)$, and $\psi_i \sim U(\phi_i, 1.1)$. For our simulations, we arbitrarily set $\lambda = 1$ (arbitrary units). Simulation parameters are summarized in Table 2. Figure 10 shows the results for 100 realizations when the number of removed nodes and initial infected vertices were set to 1% the graph size, that is, 3602 nodes. As it can be readily observed, it is more advantageous to remove nodes with higher influence power than, for exam-

Table 2 Simulation parameters—Twitter network

Parameter	Notation	Value
Number of nodes	n	360,210
Number of edges	M	14,102,583
Influence capacity	β_i^{\max}	Number of mentions + 1
Resistance to becoming infected	ϕ_i	$U(0.9, 1.1)$
Resistance to becoming a spreader	ψ_i	$U(\phi_i, 1.1)$
Resistance to attain the maximum influence power	Δ_i	100
Number of nodes removed	n_{rem}	3602

$U(a, b)$ corresponds to a uniform distribution between a and b

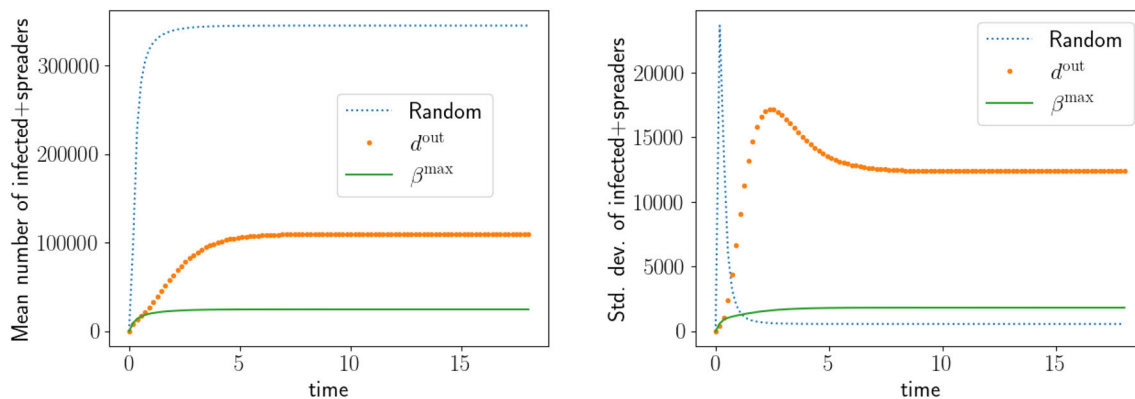


Fig. 10 Temporal evolution of the mean (left) and standard deviation of the number of infected and spreader nodes in the Twitter network

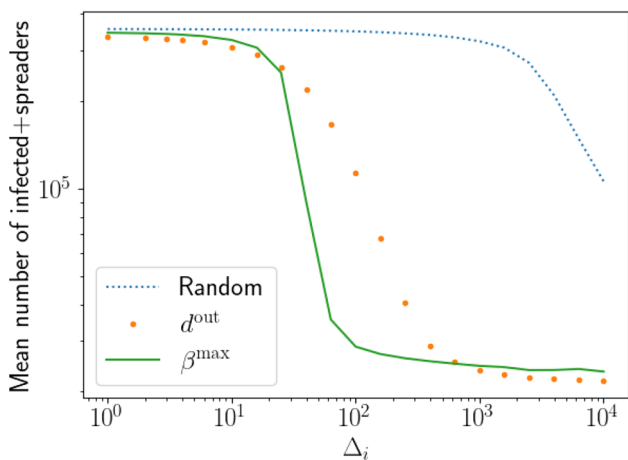


Fig. 11 Influence of Δ_i in the Twitter network. A logarithmic scale is used in order to better appreciate the differences between the three curves. The parameters ϕ_i and ψ_i were randomly assigned as $\phi_i \sim U(0.9, 1.1)$ and $\psi_i \sim U(\phi_i, 1.1)$

ple, vertices with higher degree, as we have already observed in the synthetic networks.

As we did in the previous section, we study the influence of model parameters on the results. Figure 11 shows the mean number of infected and spreader nodes at time $t = 5$ as the value of Δ_i varies. The results correspond to 100 realizations where ϕ_i and ψ_i were randomly chosen as described in the

previous paragraph. It can be observed that the strategy that removes nodes depending on β^{\max} is more convenient for intermediate values ($\Delta_i \in (30, 300)$), while not representing any real advantage over the selection using the out degree for neither smaller nor larger values. This behavior is different from that observed in the synthetic graphs (cf. right panel of Fig. 5). A plausible cause for this difference may be the positive correlation between β^{\max} and d^{out} in the Twitter network. This hypothesis is supported by the results in Fig. 8 for synthetic networks.

The right panel of Fig. 12 shows the influence of the mean value of ψ_i . The results correspond to 100 realizations, where $\Delta_i = 100$ was fixed and ϕ_i and ψ_i were varied in each simulation according to $\phi_i \sim U(0.9, 1.1)$, $\psi_i \sim U(\max(\bar{\psi} - 0.1, \phi_i), \bar{\psi} + 0.1)$. As in the synthetic networks, the benefit of using the removal strategy based on the spreading capacity diminishes as the value of $\bar{\psi}$ increases. However, a much larger change is needed in order to observe significant differences (compare the doubly logarithmic scales in the right panel of Fig. 12 to the linear ones in the right panel of Fig. 6).

The left panel of Fig. 12 exhibits the influence of the mean value of ϕ_i . However, it is difficult to disentangle its influence from that of ψ_i as it must be true that $\psi_i \geq \phi_i$. The results correspond to 100 realizations, where $\Delta_i = 100$ was fixed

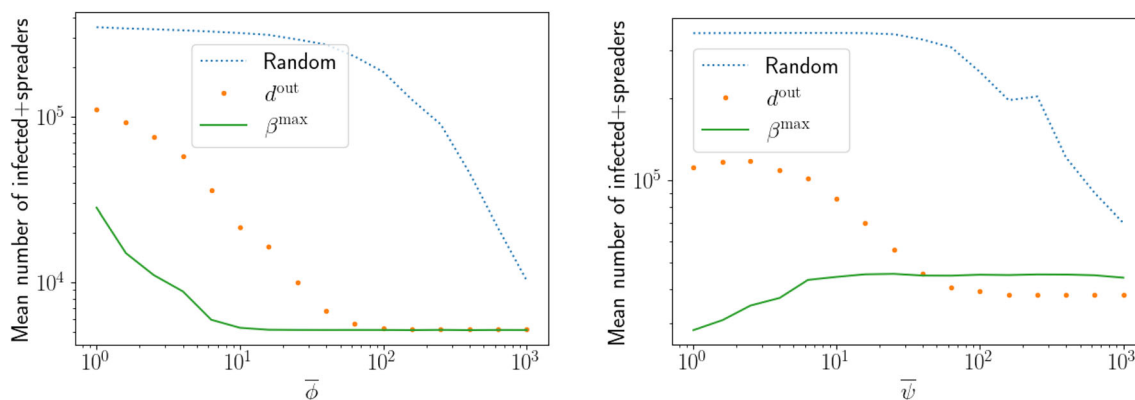


Fig. 12 Influence of $\bar{\phi}$ (left) and $\bar{\psi}$ (right) in the Twitter network. While $\Delta_i = 100$ was kept constant, ϕ_i and ψ_i were varied in all realizations

and ϕ_i and ψ_i were varied in each simulation according to $\phi_i \sim U(\bar{\phi} - 0.1, \bar{\phi} + 0.1)$, $\psi_i \sim U(\bar{\psi} - 0.1, \bar{\psi} + 0.1)$.

5 Conclusions

We studied the problem of mitigating the propagation of misinformation in networks. While most of the literature focuses on the characterization of influential agents based only on the topological properties related to their position in the social network, we introduced a new characteristic which is an intrinsic attribute of the node: its influence power.

To evaluate the impact of the influence power on the spread of fake news in the network, we put forth a new model of information diffusion. In this model, we related this attribute to an inherent potential of the agent and the intensity of its commitment to the news propagation which is, in turn, due to the strength with which such information was received. By means of numerical simulations on both synthetic and real-world networks, we showed that, in certain conditions, the influence power of the agent can be more important than its topological properties.

Our results suggest that the removal of nodes with high influence power is more effective in combating the misinformation diffusion in denser networks. Moreover, this method of mitigating fake news spread seems also to be more robust since it obtained the smallest standard deviation of the results. That strategy is also more convenient when the influence power of a few nodes is much larger than that of the general population. On the other hand, for the method to be effective it is necessary to remove a large portion of such super-spreaders.

As an important real-world example, we presented simulations based on a large Twitter dataset (> 300 thousand nodes). We posited that the number of times that a Twitter account is mentioned may represent its actual spreading capacity. Building on this idea, we showed that removing

the nodes with higher influence power is more convenient than removing those with higher degree in order to curtail the propagation of a news.

We must remark that our main objective is to demonstrate that there can exist node attributes that are more important than other structural characteristics, such as their degree, for the propagation of information. Since the main application of our results is the obstruction of viral propagation of fake news or misinformation by selective removal of agents from a social network, we believe to have made a case for accounting for the node attributes in choosing which nodes to remove.

Finally, we call the attention to certain practical difficulties related to the removal of nodes from a social network. Indeed, the subject of government policies and legal actions in relation to misinformation is a delicate matter which is very actively discussed worldwide (see, e.g., [97–102] and references therein). Banning a user from a social network can be considered an act contrary to fundamental right of free speech. Moreover, legislation can be inadequately used by totalitarian governments or ill-intentioned corporations. All in all, the methodology of removing nodes with high influence capacity, if this capacity is appropriately defined and quantified, may be more transparent for lay people than other automatic or algorithmic approaches and, hence, more desirable [101].

Author Contributions P.I.F. and L.C.R. equally contributed to the development of the ideas and wrote the main manuscript text. Data processing and simulations were the responsibility of P.I.F.

Funding This work was funded by CNPq Grant 308980/2021-2.

Declarations

Conflict of interest The authors have no conflict of interest to declare that is relevant to the content of this article.

References

1. Flat Earth Society: The Flat Earth Society website. <https://www.theflatearthsociety.org/>. Accessed Feb 2024 (2024)
2. Von Däniken, E.: *Chariots of the Gods*. Penguin, New York (1999)
3. BBC News: Flat Earth: How did YouTube help spread a conspiracy theory? <https://www.bbc.com/reel/video/p07h3yc0/flat-earth-how-did-youtube-help-spread-a-conspiracy-theory->. Accessed Feb 2024 (2019)
4. Melo, L.W.S., Passos, M.M., Salvi, R.F.: Analysis of ‘flat-earther’ posts on social media: Reflections for science education from the discursive perspective of Foucault. *Rev. Bras. Pesquisa Educ Ciências* **20**, 295–313 (2020)
5. Galhardi, C.P., Freire, N.P., Minayo, M.C.S., Fagundes, M.C.M.: Fato ou fake? Uma análise da desinformação frente à pandemia da covid-19 no Brasil. *Ciência & Saúde Coletiva* **25**, 4201–4210 (2020)
6. Jamieson, K.H.: How to debunk misinformation about covid, vaccines and masks. *Sci. Am.* **324**(4), 44–51 (2021)
7. BBC News: The vaccine misinformation battle raging in France. <https://www.bbc.com/news/blogs-trending-56526265>. Accessed Feb 2024 (2021)
8. Hotez, P., Batista, C., Ergonul, O., Figueroa, J.P., Gilbert, S., Gursel, M., Hassanain, M., Kang, G., Kim, J.H., Lall, B., et al.: Correcting Covid-19 vaccine misinformation: Lancet commission on Covid-19 vaccines and therapeutics task force members. *EClinicalMedicine* **33**, 100780 (2021)
9. World Health Organization: How to report misinformation online. <https://www.who.int/campaigns/connecting-the-world-to-combat-coronavirus/how-to-report-misinformation-online>. Accessed February 2024 (2024)
10. Cardoso, T.: Desinformação sobre vacina da COVID-19 aumenta com o início da imunização. <https://jornal.usp.br/universidade/desinformacao-sobre-vacina-da-covid-19-aumenta-com-o-inicio-da-imunizacao/>. Accessed Feb 2024 (2021)
11. Facebook for the media: Working to Stop Misinformation and False News. <https://www.facebook.com/formedia/blog/working-to-stop-misinformation-and-false-news>. Accessed Feb 2024 (2017)
12. Iosifidis, P., Nicoli, N.: The battle to end fake news: a qualitative content analysis of Facebook announcements on how it combats disinformation. *Int. Commun. Gaz.* **82**(1), 60–81 (2020)
13. Mantzarlis, A.: COVID-19: \$6.5 million to help fight coronavirus misinformation. <https://blog.google/outreach-initiatives/google-news-initiative/covid-19-65-million-help-fight-coronavirus-misinformation/>. Accessed Feb 2024 (2020)
14. Figueira, Á., Oliveira, L.: The current state of fake news: challenges and opportunities. *Procedia Comput. Sci.* **121**, 817–825 (2017)
15. Tschitschek, S., Singla, A., Gomez Rodriguez, M., Merchant, A., Krause, A.: Fake news detection in social networks via crowd signals. In: *Companion Proceedings of the The Web Conference 2018*, pp. 517–524 (2018)
16. David, L., Baum, M., Benkler, Y., Berinsky, A., Greenhill, K., Menczer, F., Metzger, M., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S.A., Sunstein, C.R., Thorson, E.A., Watts, D.J., Zittrain, J.L.: The science of fake news: addressing fake news requires a multidisciplinary effort. *Science* **359**(8), 1094–1096 (2018)
17. Zhou, X., Zafarani, R., Shu, K., Liu, H.: Fake news: Fundamental theories, detection strategies and challenges. In: *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*, pp. 836–837 (2019)
18. Twitter Inc.: Permanent suspension of @realDonaldTrump. https://blog.twitter.com/en_us/topics/company/2020/suspension.html. Accessed Feb 2024 (2021)
19. Dwoskin, E., Tiku, N.: How Twitter, on the front lines of history, finally decided to ban Trump. <https://www.washingtonpost.com/technology/2021/01/16/how-twitter-banned-trump/>. Accessed Feb 2024 (2021)
20. Applebaum, A.: Trump Is a Super-Spreader of Disinformation. <https://www.theatlantic.com/ideas/archive/2020/10/trump-super-spreader-disinformation/616604/>. Accessed Feb 2024 (2020)
21. Paul, C., Matthews, M.: *The Russian “Firehose of Falsehood” Propaganda Model: Why It Might Work and Options to Counter It*. RAND Corporation, Santa Monica, CA (2016)
22. Becket, L.: Misinformation ‘superspreaders’: COVID vaccine falsehoods still thriving on Facebook and Instagram. <https://www.theguardian.com/world/2021/jan/06/facebook-instagram-urged-fight-deluge-anti-covid-vaccine-falsehoods>. Accessed Feb 2024 (2021)
23. Sinclair, H.C.: 7 ways to avoid becoming a misinformation superspreader. <https://ideas.ted.com/7-ways-to-avoid-becoming-a-misinformation-superspreader/>. Accessed Feb 2024 (2021)
24. Watts, D.J., Dodds, P.S.: Influentials, networks, and public opinion formation. *J. Consum. Res.* **34**(4), 441–458 (2007)
25. Bakshy, E., Hofman, J.M., Mason, W.A., Watts, D.J.: Everyone’s an influencer: quantifying influence on Twitter. In: *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*, pp. 65–74 (2011)
26. Hearn, A., Schoenhoff, S.: From celebrity to influencer. In: Marshall, P.D., Redmond, S. (eds.) *A Companion to Celebrity*, pp. 194–212. Wiley, Malden (2016). (Chap. 11)
27. Woods, S.: # sponsored: The emergence of influencer marketing. Technical report, TRACE: Tennessee Research and Creative Exchange, University of Tennessee (2016)
28. De Veirman, M., Cauberghe, V., Hudders, L.: Marketing through instagram influencers: the impact of number of followers and product divergence on brand attitude. *Int. J. Advert.* **36**(5), 798–828 (2017)
29. Jin, S.V., Muqaddam, A., Ryu, E.: Instafamous and social media influencer marketing. *Mark. Intell. Plan.* **37**, 567–579 (2019)
30. Kitsak, M., Gallos, L.K., Havlin, S., Liljeros, F., Muchnik, L., Stanley, H.E., Makse, H.A.: Identification of influential spreaders in complex networks. *Nat. Phys.* **6**(11), 888–893 (2010)
31. Pei, S., Makse, H.A.: Spreading dynamics in complex networks. *J. Stat. Mech: Theory Exp.* **2013**(12), 12002 (2013)
32. Pei, S., Muchnik, L., Andrade, J.S., Jr., Zheng, Z., Makse, H.A.: Searching for superspreaders of information in real-world social media. *Sci. Rep.* **4**(1), 1–12 (2014)
33. Taha, K.: Identifying the top-k influential spreaders in social networks: a survey and experimental evaluation. *IEEE Access* **10**, 107809–107845 (2022)
34. Andrade, R.L., Rêgo, L.C., Silva, T.L.C., Macêdo, J.A.F., Silva, W.C.: Energy disruptive centrality with an application to criminal network. *Commun. Nonlinear Sci. Numer. Simul.* **99**, 105834 (2021)
35. Hegselmann, R., Krause, U., et al.: Opinion dynamics and bounded confidence models, analysis, and simulation. *J. Artif. Soc. Soc. Simul.* **5**(3), 1–33 (2002)
36. Kuperman, M., Zanette, D.: Stochastic resonance in a model of opinion formation on small-world networks. *Eur. Phys. J. B-Condens. Matter Complex Syst.* **26**(3), 387–391 (2002)
37. Tesson, C.J., Toral, R.: System size stochastic resonance in a model for opinion formation. *Physica A* **351**(1), 106–116 (2005)

38. Holme, P., Newman, M.E.: Nonequilibrium phase transition in the coevolution of networks and opinions. *Phys. Rev. E* **74**(5), 056108 (2006)
39. Li, P.-P., Zheng, D.-F., Hui, P.: Dynamics of opinion formation in a small-world network. *Phys. Rev. E* **73**(5), 056128 (2006)
40. Nardini, C., Kozma, B., Barrat, A.: Who's talking first? Consensus or lack thereof in coevolving opinion formation models. *Phys. Rev. Lett.* **100**(15), 158701 (2008)
41. Tessone, C.J., Toral, R.: Diversity-induced resonance in a model for opinion formation. *Eur. Phys. J. B* **71**(4), 549–555 (2009)
42. Jalili, M.: Social power and opinion formation in complex networks. *Physica A* **392**(4), 959–966 (2013)
43. Acemoğlu, D., Como, G., Fagnani, F., Ozdaglar, A.: Opinion fluctuations and disagreement in social networks. *Math. Oper. Res.* **38**(1), 1–27 (2013)
44. Gionis, A., Terzi, E., Tsaparas, P.: Opinion maximization in social networks. In: *Proceedings of the 2013 SIAM International Conference on Data Mining*, pp. 387–395 (2013). SIAM
45. Sirbu, A., Loreto, V., Servidio, V.D., Tria, F.: Opinion dynamics: models, extensions and external effects. In: Loreto, V., Haklay, M., Hotho, A., Servidio, V.D.P., Stumme, G., Theunis, J., Tria, F. (eds.) *Participatory Sensing, Opinions and Collective Awareness*, pp. 363–401. Springer, Cham (2017)
46. Medo, M., Mariani, M.S., Lü, L.: The fragility of opinion formation in a complex world. *Commun. Phys.* **4**(1), 1–10 (2021)
47. Hethcote, H.W.: The mathematics of infectious diseases. *SIAM Rev.* **42**(4), 599–653 (2000)
48. Pastor-Satorras, R., Castellano, C., Van Mieghem, P., Vespignani, A.: Epidemic processes in complex networks. *Rev. Mod. Phys.* **87**(3), 925 (2015)
49. Katz, E.: The two-step flow of communication: An up-to-date report on an hypothesis. *Public Opin. Q.* **21**(1), 61–78 (1957)
50. Nisbet, M.C., Kotcher, J.E.: A two-step flow of influence? Opinion-leader campaigns on climate change. *Sci. Commun.* **30**(3), 328–354 (2009)
51. Weimann, G.: The influentials: back to the concept of opinion leaders? *Public Opin. Q.* **55**(2), 267–279 (1991)
52. Trepte, S., Scherer, H.: Opinion leaders-do they know more than others about their area of interest? *Communications* **35**(2), 119–140 (2010)
53. Page, B.I., Shapiro, R.: Presidents as opinion leaders: Some new evidence. *Policy Stud. J.* **12**(4), 649 (1984)
54. Venkatraman, M.P.: Opinion leaders, adopters, and communicative adopters: a role analysis. *Psychol. Mark.* **6**(1), 51–68 (1989)
55. Park, C.S.: Does twitter motivate involvement in politics? tweeting, opinion leadership, and political engagement. *Comput. Hum. Behav.* **29**(4), 1641–1648 (2013)
56. Winter, S., Neubaum, G.: Examining characteristics of opinion leaders in social media: a motivational approach. *Soc. Med. Soc.* **2**(3), 2056305116665858 (2016)
57. Valente, T.W., Pumpuang, P.: Identifying opinion leaders to promote behavior change. *Health Educ. Behav.* **34**(6), 881–896 (2007)
58. Bamakan, S.M.H., Nurgaliev, I., Qu, Q.: Opinion leader detection: a methodological review. *Expert Syst. Appl.* **115**, 200–222 (2019)
59. Nie, Y., Li, W., Pan, L., Lin, T., Wang, W.: Markovian approach to tackle competing pathogens in simplicial complex. *Appl. Math. Comput.* **417**, 126773 (2022)
60. Nie, Y., Zhong, X., Lin, T., Wang, W.: Homophily in competing behavior spreading among the heterogeneous population with higher-order interactions. *Appl. Math. Comput.* **432**, 127380 (2022)
61. Newman, M.: *Networks*. Oxford University Press, Oxford (2018)
62. Bonacich, P.: Some unique properties of eigenvector centrality. *Soc. Netw.* **29**(4), 555–564 (2007)
63. Brin, S., Page, L.: The anatomy of a large-scale hypertextual web search engine. *Comput. Netw. ISDN Syst.* **30**(1–7), 107–117 (1998)
64. Anthonisse, J.M.: The rush in a directed graph. Technical Report BN 9/71, Stichting Mathematisch Centrum. *Mathematische Besliskunde* (1971)
65. Freeman, L.C.: Centrality in social networks. *Conceptual clarification. Soc. Netw.* **1**(3), 215–239 (1978)
66. Xiong, F., Liu, Y., Zhang, Z.-J., Zhu, J., Zhang, Y.: An information diffusion model based on retweeting mechanism for online social media. *Phys. Lett. A* **376**(30–31), 2103–2108 (2012)
67. Melnik, S., Ward, J.A., Gleeson, J.P., Porter, M.A.: Multi-stage complex contagions. *Chaos* **23**(1), 013124 (2013)
68. Huang, W.-M., Zhang, L.-J., Xu, X.-J., Fu, X.: Contagion on complex networks with persuasion. *Sci. Rep.* **6**(1), 1–8 (2016)
69. Centola, D., Macy, M.: Complex contagions and the weakness of long ties. *Am. J. Sociol.* **113**(3), 702–734 (2007)
70. Bakshy, E., Rosenn, I., Marlow, C., Adamic, L.: The role of social networks in information diffusion. In: *Proceedings of the 21st International Conference on World Wide Web*, pp. 519–528 (2012)
71. Gleeson, J.P.: Cascades on correlated and modular random networks. *Phys. Rev. E* **77**(4), 046117 (2008)
72. Dodds, P.S., Payne, J.L.: Analysis of a threshold model of social contagion on degree-correlated networks. *Phys. Rev. E* **79**(6), 066115 (2009)
73. Gai, P., Kapadia, S.: Contagion in financial networks. *Proc. R. Soc. A: Math. Phys. Eng. Sci.* **466**(2120), 2401–2423 (2010)
74. Romero, D.M., Meeder, B., Kleinberg, J.: Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on Twitter. In: *Proceedings of the 20th International Conference on World Wide Web*, pp. 695–704 (2011)
75. Barash, V., Cameron, C., Macy, M.: Critical phenomena in complex contagions. *Soc. Netw.* **34**(4), 451–461 (2012)
76. Dodds, P.S., Harris, K.D., Danforth, C.M.: Limited imitation contagion on random networks: chaos, universality, and unpredictability. *Phys. Rev. Lett.* **110**(15), 158701 (2013)
77. Karimi, F., Holme, P.: Threshold model of cascades in empirical temporal networks. *Physica A* **392**(16), 3476–3483 (2013)
78. Wang, W., Tang, M., Zhang, H.-F., Lai, Y.-C.: Dynamics of social contagions with memory of nonredundant information. *Phys. Rev. E* **92**(1), 012820 (2015)
79. González-Bailón, S., Borge-Holthoefer, J., Rivero, A., Moreno, Y.: The dynamics of protest recruitment through an online network. *Sci. Rep.* **1**(1), 1–7 (2011)
80. Centola, D.: An experimental study of homophily in the adoption of health behavior. *Science* **334**(6060), 1269–1272 (2011)
81. Centola, D.M.: Homophily, networks, and critical mass: solving the start-up problem in large group collective action. *Ration. Soc.* **25**(1), 3–40 (2013)
82. Kim, D.A., Hwang, A.R., Stafford, D., Hughes, D.A., O'Malley, A.J., Fowler, J.H., Christakis, N.A.: Social network targeting to maximise population behaviour change: a cluster randomised controlled trial. *Lancet* **386**(9989), 145–153 (2015)
83. DellaPosta, D., Nee, V., Opper, S.: Endogenous dynamics of institutional change. *Ration. Soc.* **29**(1), 5–48 (2017)
84. Aral, S., Dhillon, P.S.: Social influence maximization under empirical influence models. *Nat. Hum. Behav.* **2**(6), 375–382 (2018)
85. Guilbeault, D., Centola, D.: Topological measures for identifying and predicting the spread of complex contagions. *Nat. Commun.* **12**(1), 1–9 (2021)
86. Aral, S., Muchnik, L., Sundararajan, A.: Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proc. Natl. Acad. Sci.* **106**(51), 21544–21549 (2009)

87. Centola, D.: Failure in complex social networks. *J. Math. Sociol.* **33**(1), 64–68 (2008)
88. Kuhlman, C.J., Anil Kumar, V., Marathe, M.V., Ravi, S., Rosenkrantz, D.J.: Inhibiting diffusion of complex contagions in social networks: theoretical and experimental results. *Data Min. Knowl. Disc.* **29**(2), 423–465 (2015)
89. Carscadden, H.L., Kuhlman, C.J., Marathe, M.V., Ravi, S., Rosenkrantz, D.J.: Blocking the propagation of two simultaneous contagions over networks. In: *International Conference on Complex Networks and Their Applications*, pp. 455–468 (2020). Springer
90. Lalou, M., Tahraoui, M.A., Kheddouci, H.: The critical node detection problem in networks: a survey. *Comput. Sci. Rev.* **28**, 92–117 (2018)
91. Cavallaro, L., Costantini, S., De Meo, P., Liotta, A., Stilo, G.: Network connectivity under a probabilistic node failure model. *IEEE Trans. Netw. Sci. Eng.* **9**(4), 2463–2480 (2022)
92. Myers, S.A., Zhu, C., Leskovec, J.: Information diffusion and external influence in networks. In: *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 33–41 (2012)
93. Gillespie, D.T.: Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **81**(25), 2340–2361 (1977)
94. Barabási, A.-L., Albert, R.: Emergence of scaling in random networks. *Science* **286**(5439), 509–512 (1999)
95. Bakshy, E., Karrer, B., Adamic, L.A.: Social influence and the diffusion of user-created content. In: *Proceedings of the 10th ACM Conference on Electronic Commerce*, pp. 325–334 (2009)
96. De Domenico, M., Lima, A., Mougél, P., Musolesi, M.: The anatomy of a scientific rumor. *Sci. Rep.* **3**(1), 1–9 (2013)
97. Bayer, J., Bitiukova, N., Bard, P., Szakács, J., Alemanno, A., Uszkiewicz, E.: Disinformation and propaganda-impact on the functioning of the rule of law in the eu and its member states. European Parliament, LIBE Committee, Policy Department for Citizens' Rights and Constitutional Affairs (2019)
98. Bayer, J., Holznagel, D.B., Katarzyna, L., Adela, P., Josephine, B.S., Szakács, J., Uszkiewicz, E.: Disinformation and propaganda: impact on the functioning of the rule of law and democratic processes in the eu and its member states: 2021 update (2021)
99. Bontcheva, K., Posetti, J. (eds.): *Balancing Act: Countering Digital Disinformation While Respecting Freedom of Expression*. Broadband Commission Research Report on 'Freedom of Expression and Addressing Disinformation on the Internet'. ITU, UNESCO, Geneva and Paris (2020)
100. Marsden, C., Meyer, T., Brown, I.: Platform values and democratic elections: how can the law regulate digital disinformation? *Comput. Law Secur. Rev.* **36**, 105373 (2020)
101. Kyza, E.A., Varda, C., Panos, D., Karageorgiou, M., Komendantova-Amann, N., Perfumi, S.C., Shah, S.I.H., Hosseini, A.S.: Combating misinformation online: re-imagining social media for policy-making. *Internet Policy Rev.* **9**(4), 1–24 (2020)
102. Ó Fathaigh, R., Helberger, N., Appelman, N.: The perils of legally defining disinformation. *Amsterdam Law School Research Paper*, 2022-05 (2021)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.