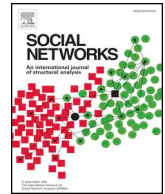




ELSEVIER

Contents lists available at ScienceDirect

Social Networks

journal homepage: www.elsevier.com/locate/socnet

Community detection in civil society online networks: Theoretical guide and empirical assessment

Daniela Stoltenberg^{a,*}, Daniel Maier^b, Annie Waldherr^a^a Department of Communication, University of Münster, Germany^b Institute for Media and Communication Studies, Free University Berlin, Germany

ARTICLE INFO

Keywords:

Hyperlink network analysis
Community detection
Cohesive subgroup
Civil society
Computational methods

ABSTRACT

Community detection is a fundamental challenge in the analysis of online networks. However, there is a lack of consensus regarding how to accomplish this task in a manner that acknowledges domain-specific, substantive social theory. We develop a typology of what social phenomena communities of hyperlinked actors may signify—topical similarities, ideological associations, strategic alliances, and potential user traffic—and offer recommendations for community detection grounded in these concepts. Testing procedures on a hyperlink network of the food safety movement, we demonstrate that the handling of tie directions and weights as well as algorithm choice influence which communities are ultimately detected in such a network.

Over the past decade, the study of networks has rapidly gained traction among social scientists investigating (online) civil society (e.g., Bennett and Segerberg, 2013; Castells, 2010). Many of these studies collect hyperlink networks—that is, networks in which websites represent nodes and hyperlinks the directed ties among them (Park, 2003). These networks are often studied as “issue networks” in the sense of hosting public debates on specific issues such as climate change (Rogers and Marres, 2000), fair trade (Bennett et al., 2011), or online copyright law (Benkler et al., 2015).

Furthermore, hyperlink networks are analyzed to infer meaningful social relationships, as hyperlinks may signify collective identities, discursive coalitions, or ideological affinity among actors (Sereno, 2010). Hyperlinks also offer resources to civil society organizations trying to mobilize for their interests. First, they grant *visibility*. They structure the distribution of attention across the Internet by directing users to websites (Barzilai-Nahon, 2008). Second, links act as recommendations signifying common interests and affiliations (González-Bailón, 2009), which lend *prestige* to a website. Even in the age of social media, the relevance of these functions of hyperlinks remains. Websites are still the public faces of organizations (Bennett and Segerberg, 2013), and social actors use social media to connect followers and sympathizers to the content on their website (Nitschke et al., 2014). Furthermore, hyperlinks are crucial for the importance that is attributed to websites by search engine algorithms (Brin and Page, 1998).

Since hyperlinks are not uniformly distributed, cohesive node clusters or communities emerge in networks (González-Bailón, 2009).

These are “subsets of actors among whom there are relatively strong, direct, intense, frequent or positive ties” (Wasserman and Faust, 2009, p. 249). From the perspective of resource mobilization theory (McCarthy and Zald, 1977), such communities represent opportunity structures for civil society actors (Kriesi et al., 1992). They facilitate the spread of collective action frames (Snow et al., 1986) and the mobilization of resources such as attention, prestige, or information (Baldassarri and Diani, 2007). In the growth phases of social movements, network communities are relevant for creating a critical mass of like-minded, interconnected actors (Centola, 2013).

Consequently, there is an increasing usage of community detection procedures by researchers working on online civil society networks. However, the range of community detection algorithms is wide, as is the range of interpretations of what communities are, often based on specific theoretical concepts from sociology and public spheres research. Some researchers interpret communities as topical clusters of larger issue networks in the sense of *issue-specific sub-publics* (e.g., Ackland and O’Neil, 2011; Herring et al., 2005). Others add ideological homophily in terms of issue positions, which allows them to differentiate between *publics* and *counter-publics* (Benkler et al., 2015; Kaiser and Puschmann, 2017). To the extent that different actors within communities share the same beliefs and stories about issues, communities can also be interpreted as *discourse coalitions*, as introduced by Hajer (1995). An even stronger collective identity is assumed by researchers who define communities as strategic alliances (Pilny and Shumate, 2012). Referencing Sabatier (1988), Adam et al. (2018) term

* Corresponding author at: Garystr. 55, 14195 Berlin, Germany.

E-mail address: daniela.stoltenberg@uni-muenster.de (D. Stoltenberg).

<https://doi.org/10.1016/j.socnet.2019.07.001>

Available online 19 July 2019

0378-8733/ © 2019 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

these alliances *advocacy coalitions* if they are formed in response to specific policy issues, share the same beliefs, show a substantial degree of coordination, and are stable over time. When referring to theory or theoretical considerations in the following, we mean this type of *domain-specific* or *substantive* theory.

On the methodological side, many community detection algorithms exist, and researchers have to make decisions concerning the preprocessing of networks. For instance, not all algorithms are capable of handling the natural features of hyperlinks—their directedness and weight. Therefore, either hyperlink data has to be preprocessed or domain-specific theoretical considerations may support the inclusion of tie weights and directions, thereby restricting the choice of algorithm. Although the goal of community detection is identifying natural structures within a network (Newman, 2006), we show that decisions regarding tie weights and directionality, as well as the choice of a community detection algorithm, critically influence the uncovered community structure.

As Peel et al. (2017) have argued, the diversity of options is not a weakness, as many mechanisms may have contributed to a network's formation and not all procedures will be equally appropriate for uncovering each process. Accordingly, Ghasemian et al. (2018) found that results of different algorithms vary widely for the same data and that for different circumstances, different algorithms perform best.

Existing studies comparing different community detection algorithms mostly evaluate algorithms on artificial networks with respect to a known ground truth (Fortunato, 2010; Papadopoulos et al., 2012; Yang et al., 2016) or on empirical networks with respect to performance-based benchmarks (Ghasemian et al., 2018; Leskovec et al., 2010). We argue that the decisions required in community detection should also be based on domain-specific theoretical considerations. This is important for two reasons: First, performance-based assessment alone does not help in choosing a specific methodological design (as several algorithms might perform equally well with respect to general benchmarks); second, when studying large online networks of civil society actors there is often no clear ground truth against which performance might be tested.

The main contribution of this paper is to provide a heuristic scheme that matches the theoretically derived social meanings of hyperlinks among websites of civil society actors with the most appropriate methods for finding communities within such networks. Our focus is on public/political debate in issue networks, as this is one of the most intensively studied areas in hyperlink network analysis (De Maeyer, 2013).

The remainder of the article is organized as follows. First, we briefly introduce the challenge of community detection. We then synthesize and discuss the research literature, focusing on two questions: What kind of social structures may communities in hyperlink networks represent? And which methods are most appropriate for which kind of community? We conclude by offering guidelines on how to choose the method of community detection based on the specific aims of the research. Second, we evaluate different approaches to community detection for an empirical hyperlink network of the food safety movement in the United Kingdom. In offering a theoretically grounded rationale for the decisions required for community detection and discussing their consequences, we provide guidance to researchers in the field of communication networks online.

Theoretical foundations of community detection

Decisions on community detection for hyperlink networks

Communities or cohesive subgroups are “groups of actors who interact with each other to such an extent that they could be considered to be a separate entity” (Borgatti et al., 2013, p. 181). Because several properties of subgroups, such as their relative relational density and degree of connectivity, may be considered relevant (Barabási, 2016), no

single operational definition of communities exists. Here, we discuss three decisions relevant in community detection: choice of algorithm, handling of tie directions and tie weights.

Algorithms

Because hyperlink networks tend to contain at least a few hundred nodes, there are some constraints on community detection methods. Concepts defining clear-cut criteria for intra-group cohesion, such as n -cliques or k -plexes (Wasserman and Faust, 2009), may not yield useful results, identifying either very few or very many largely overlapping subgroups and frequently leaving many nodes unassigned. Therefore, algorithmic procedures have become the most applied strategies. In contrast to previous deductive approaches, they specify the process of community detection, rather than the communities themselves (Borgatti et al., 2013). A multitude of such procedures exist (see Fortunato, 2010 for an overview), and many are readily implemented in software. In the following, for reasons of applicability, accessibility, and practical relevance, we will focus on five algorithms that are implemented in the *igraph* package (Csardi and Nepusz, 2006) for the statistical programming environment R (R Development Core Team, 2008) and that are widely used in social scientific research. A comprehensive review of community detection procedures exceeds the scope of this article (for more extensive discussion, see e.g., Fortunato, 2010; Papadopoulos et al., 2012).

The five algorithms are each grounded on different principles. The *map equation* algorithm (Rosvall et al., 2009) is based on the movement of a random walker through the network. Its aim is to optimize a global two-level code in such a way that nodes the walker frequently visits in sequence are grouped together. The algorithm is based on an agglomerative principle. It starts out with each node as its own community and subsequently merges them. The *label propagation* algorithm (Raghavan et al., 2007) assigns temporary labels to each node and then investigates the communities that emerge when the nodes iteratively adopt their nearest neighbors' labels. Therefore, this algorithm is also agglomerative; however, it is based on the emergence of local properties, matching the intuitive notion of group structure. Both of the above algorithms contain stochastic processes. Accordingly, they will not discover completely identical communities when applied multiple times.

The other algorithms, *edge betweenness* (Girvan and Newman, 2002; Newman and Girvan, 2004), *leading eigenvector* (Newman, 2006), and *multilevel* (Blondel et al., 2008), are each based on the optimization of the partitions' modularity. This metric describes the proportion of intra-community ties, compared to the one expected in a random network (Newman and Girvan, 2004). The *edge betweenness* algorithm focuses on identifying the ties with the highest betweenness, that is, “the [highest] number of shortest paths between pairs of vertices that run along it” (Girvan and Newman, 2002, p. 7822). Edges with a high betweenness will likely run between communities, where there are only relatively few connections. The algorithm operates divisively by iteratively eliminating the ties with the highest betweenness so that the network begins to separate into communities. Modularity is used to determine when the procedure has achieved an optimum community structure.

The *leading eigenvector* algorithm employs modularity by identifying the modularity matrix's leading eigenvector and sequentially bisecting the network accordingly. After each division, the leading eigenvector is calculated for the modularity matrix of each component and the process repeats iteratively until further bisections will not yield an improvement in modularity. Finally, the *multilevel* algorithm initially treats all nodes as individual communities and successively moves nodes between communities. This agglomerative process also uses modularity to determine the optimal community solution.

Comparative studies suggest that results vary significantly across different algorithms. One strand of research focuses on performance evaluation on artificial benchmark graphs. Yang et al. (2016) evaluated the properties of several algorithms on artificial benchmark graphs by

varying the mixing parameter, that is, the share of inter-community ties. If the mixing parameter was small, all algorithms performed comparatively well. However, most algorithms exhibited a turning point value beyond which performance decreased rapidly. Papadopoulos et al. (2012) focused on performance in terms of computational complexity and memory requirements and found that procedures showed marked differences, with label propagation and multilevel being comparatively computationally cheap and therefore applicable even to very large networks.

Other work focuses on empirical networks. In a comparison of 16 different procedures on a large corpus of empirical networks, Ghasebian et al. (2018) showed that outputs differed widely in the number and composition of communities. They argue that different algorithms do not outperform others in general, but performance is based on the specific network. Similarly, Leskovec et al. (2010) showed that results for different community detection procedures vary both in their quality (i.e., how well procedures perform in approaching an optimal group structure) and the structural properties of their solutions.

These studies indicate that different algorithms come to rather divergent results, depending on the properties of the network. However, researchers may be left wondering what method to choose and why (Hartman et al., 2017). We argue that integrating the notion of what social phenomena communities indicate is necessary to this end. Hartman et al. (2017) also make this proposition and introduce a rank stability measure to evaluate how well a community structure aligns with ground truth metadata. They applied this to Reddit user interaction data and tested how well community structure aligns with common SubReddits.

This is a step towards connecting network structure and the social meaning of nodal relations, as we propose. However, we take it a step further: Rather than proposing an ex-post evaluation based on a known ground truth, we highlight the necessity for domain-specific theoretical considerations. Grounding selection of community detection algorithms and respective preprocessing steps in sociological concepts (in our case, referring to theories of the public sphere, civil society, and social movements) ensures construct validity of the method. This is particularly relevant if a pre-evaluated community ground truth is unavailable, as is often the case when exploring community structures of online networks.

Tie weights and directions

Hyperlink ties are generally directed (one website links to another) and weighted (more than one hyperlink may exist between two websites). Due to the different capabilities of community detection algorithms, however, the directedness and weightedness of edges are frequently modified by symmetrizing the adjacency matrix and dichotomizing weights. We focus on tie directions and weights here because, while other graph features, such as density, degree distribution, and diameter (Leskovec et al., 2010) have consequences for the resulting graph partitioning too, they are not the target of transformation. Moreover, most hyperlink networks exhibit similar features, i.e., a sparsity of ties, long-tail degree distributions and short diameters.

A review of research on hyperlink networks shows that the directionality of ties is frequently ignored, while tie weights are usually retained (e.g., Ackland and O’Neil, 2011; Pfetsch et al., 2016). However, this choice and its implications are not frequently discussed. We argue that, depending on the research question, either retaining or ignoring tie weights may be reasonable. The latter option requires the definition of a threshold value w , a minimum weight required for ties to be retained. For $w = 1$, all ties remain, whereas for $w = 2$, ties with $w < 2$ are ignored while all remaining ties are treated equally.

All algorithms discussed here are capable of handling weighted and unweighted networks, but in their implementation in the *igraph* package (Csardi and Nepusz, 2006), most cannot handle directed ties. Only the *map equation* and the *edge betweenness* algorithms detect communities in directed networks. To circumvent this problem, one can symmetrize the network’s adjacency matrix either by keeping reciprocated ties only or by ignoring the direction of the ties. Most hyperlink networks are sparse (Barabási, 2016) and have few reciprocated hyperlinks. Thus, retaining only reciprocated ties will dramatically shrink the number of connected actors.

The social meaning of communities in hyperlink networks

Social scientists are usually not interested in hyperlinks as purely technical artifacts. Instead, the underlying assumption is that hyperlinks signify meaningful social relations. Regarding hyperlinked communities, the hypothesis is that “members’ differ from ‘non-members’ in theoretically important ways” (Wasserman and Faust, 2009, p. 283). To define what constitutes a meaningful difference, researchers need to determine what features would constitute a community based on the domain-specific theory guiding their work. Then, translating these features into appropriate network measures ensures construct validity.

Referring to the literature on civil society hyperlink networks, we extracted four concepts that hyperlinks may represent (see Fig. 1). In most studies, hyperlinks are used *diagnostically* to learn about the relationships between senders and receivers of links—that is, *topical similarity*, *ideological association*, or *strategic alliance*. The fourth concept is *prognostic*, as it is concerned with the consequences of hyperlink structures.

The lowest level of actor relationships is *topical similarity* (Park and Thelwall, 2003). Hyperlinks between websites may indicate topical similarity because actors establish links to provide context information, improve the informational value of a website (Park and Thelwall, 2008), or enhance the salience of embedded information (Benkler et al., 2015). Communities based on topical similarity would be characterized by the actors’ interest in the same issue, but not necessarily by a shared opinion.

Empirically, the observation of communities of actors with similar topics linking to each other on the web was made quite early (Herring et al., 2005; Kleinberg and Lawrence, 2001). Even within issue networks that focus on the same overarching topic, such as the environmental movement (Ackland and O’Neil, 2011), community detection

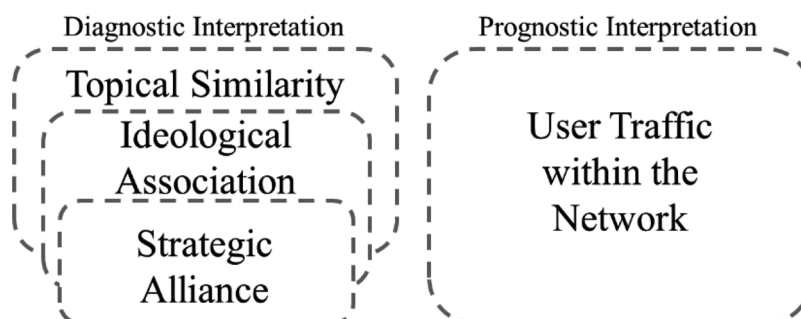


Fig. 1. Interpretations of subgroups in civil society online networks.

has proven capable of identifying communities focusing on sub-topics. Because topical similarity is a weak relationship criterion, we argue that it demands only a low level of cohesion in community detection.

The second level of actor relationships can be described as *ideological association* and is probably the most popular interpretation of hyperlinking behavior in social movement research. Accordingly, hyperlinks indicate common goals (Bennett and Segerberg, 2013; Burris et al., 2000) or a collective identity (Pily and Shumate, 2012). Ackland and O’Neil (2011) remark that “[r]eceiving hyperlinks is a form of endorsement” (p. 180). Conversely, the act of non-linking has been interpreted as boundary definition (Rogers and Marres, 2000). We can assume that actors develop some idea of belonging to the same collective, as the definition of social movements suggests (Diani, 1992). For example, Adamic (1999) discovered separate communities of pro-life and pro-choice activists in the online debate about abortion, and both Hyun (2012) and Adamic and Glance (2005) found divides between conservative and liberal bloggers in the United States. Unlike topical similarity, ideological association requires shared positions, beliefs, or narratives amongst actors forming a community. Accordingly, hyperlink communities can be interpreted as *discourse coalitions* (Hajer, 1995).

Interpreting hyperlinks as indicative of *strategic alliances* (Baldassarri and Diani, 2007) or *advocacy coalitions* (Adam et al., 2018; Sabatier, 1988) presupposes that hyperlinks provide resources, such as visibility or prestige. Actors strategically use hyperlinks to take advantage of these resources. Hyperlinking is seen as indicative of collective action (González-Bailón, 2009; Shumate and Dewitt, 2008). Rational actors form an alliance to work toward a collective good that none of them could produce on their own (Shumate and Lipp, 2008). Hyperlinking can facilitate civil society actors’ efforts by increasing issue salience as well as the prominence of the actors, enhancing mobilization and coalition building, and improving access to stakeholders (Pfetsch et al., 2013). Strategic alliances require actors’ commitment to contribute, coordinate, and communicate. Altogether, these demands translate into a high level of cohesion.

These three diagnostic interpretations may be read as hierarchical, that is, groups of strategically allied actors are likely to also be on the same ideological side of a debate and concerned with similar topics. Meanwhile, it is entirely possible for actors to be topically similar without being ideologically associated, let alone strategically allied. These differences in the degree of cohesion between actors should be reflected in the operationalization of communities. It must be acknowledged that this categorization represents ideal types. Occasionally, deviant patterns of linking have been described (Serenó, 2010).

While the previous three categories are diagnostic, the fourth category is *prognostic* in that it is concerned with the consequences of link structures—that is, user traffic. Although Wu and Ackland (2014) showed a mismatch between hyperlinks and the clickstream network inferred from user movement at the global level, this finding might not hold for country- and issue-specific networks, which on aggregate reflect user movement patterns of browsing by topic more closely (Bruno, 2007). While we cannot assume a near-perfect correlation between hyperlink networks and user traffic, the former are often the only indication of the latter that researchers can access. Consequently, links are interpreted as making information available to an audience (Benkler et al., 2015) or providing non-members with orientation within the issue network (Shumate and Dewitt, 2008). In this sense, communities represent densely interlinked areas of networks to which users are likely to remain bounded while surfing the web. Hyperlinks, thus, structure flow processes, and strong community structure points toward structural holes that may hamper this flow (González-Bailón and Wang, 2016).

Finding theory-based communities in hyperlink networks

Drawing on this typology, we argue that not all methodological options for community detection are equally appropriate in all cases.

Because hyperlinking is a cheap form of communication, the danger of overstating the social-scientific informative value of communities is high. Therefore, we need to align the methodological decisions with regards to preprocessing the network with the theoretically derived concept of community. Treating all (e.g., reciprocal and asymmetric, or strong and weak) hyperlinks equally, regardless of the type of community that researchers are seeking, might lead to erroneous conclusions. The required degree of cohesion is higher when we are searching for ideologically associated communities or strategically allied communities than for topically similar actor groups. If conclusions about user flow are drawn, different demands must once again be met.

Adamic and Glance (2005) provided a striking example of the importance of tie directions and weights. Analyzing a network of 40 US political blogs (20 liberal and 20 conservative), they found that keeping all hyperlink ties, regardless of weight and direction, resulted in a densely connected network. Gradually increasing the weight threshold and demanding reciprocity finally led to the network’s division into two components along partisan lines. This illustrates that setting a low weight threshold and disregarding the directionality of hyperlinks results in one densely connected group that is characterized by topical similarity: all blogs deal with US politics, after all. Meanwhile, altering the threshold value and demanding reciprocity leads to a divide based on ideology. What do these considerations mean for the three methodological decisions outlined above?

Tie directions

The decision on how to handle tie directions is fundamental. Because few community detection algorithms handle directed networks, researchers may be tempted to symmetrize ties. But if they want to draw conclusions on user traffic through hyperlink networks, they need to use a directed network. Community detection in the undirected (symmetrized) version of the network might lead to communities that would not allow navigation within the community.

Symmetrizing may be appropriate, if we are interested in undirected relationships—for example, if we want to learn which actors are *topically similar* or *strategically allied*. For symmetrizing, there are two options—either treat all ties equally, irrespective of directionality, or take only reciprocal ties into account. While reciprocal ties are the preferred choice in the analysis of *social* networks (Borgatti et al., 2013) such as friendship networks, they can be impractical for large, sparse hyperlink networks. As most hyperlinks are asymmetric relations, reciprocal-tie symmetrizing erases them, and the network will fall apart. We believe, however, that this can be necessary when looking for relationships beyond mere content similarity. If the aim is to identify cooperating actors, reciprocal links should be focused, because the “relation ‘is strategically allied with’ is a bidirectional relation” (Monge and Contractor, 2003, p. 35).

The order in which network data is preprocessed is important. We suggest adjusting tie directions first. Possible weight dichotomization should be executed only thereafter, because symmetrizing reciprocal ties requires summing the weights of both directions. This may affect whether the resulting tie weight is above a certain threshold.

Tie weights

Most studies retain tie weights, implying that actors who link to each other frequently are likely to be grouped together. However, researchers should consider whether a dyad sharing 30 hyperlinks should be treated as three times as close as one that shares 10 hyperlinks or if there is a threshold, beyond which the exact number of links no longer matters. From our point of view, frequent hyperlinks signify stronger relationships than single connections. Therefore, thresholding may be a viable choice to focus strong ties when investigating, for example, ideological association. However, it may be misleading to treat tie weight as directly proportional to the quality of a relationship. Therefore, researchers should critically examine whether weights add useful information for community detection. There may also be factors

introducing additional noise: for example, the heterogeneity of website architecture leads to diversity in document volumes and design features.¹ Websites with a higher volume of documents have a higher chance to link to others more frequently and are more likely to be classified as central actors in communities.

Testing different threshold values reveals the community structure at different degrees of cohesion (Borgatti et al., 2013). Herring et al. (2005) showed that when changing the threshold, some communities fall apart, while others remain stable. In any case, an increased threshold prevents an overstatement of the importance of infrequent intra-community ties when searching for strong-tie communities such as strategic alliances.²

There are also cases in which tie weights are clearly useful to retain, for example, when studying potential user movement. Here, users are more likely to click from one site to another if there are more links between the two.

Choosing an algorithm

Of the algorithms considered, only *edge betweenness* and *map equation* are capable of handling directed networks in *igraph* (Csardi and Nepusz, 2006) (see Table 1). *Map equation* may be the better choice for finding communities in which users can linger, because it assesses the possibility of flow within groups. In contrast, *edge betweenness* does not investigate intra-group paths and may create communities with limited possibilities of flow.

Beyond this restriction, it is difficult to make a priori predictions as to which procedures may be most appropriate for which research question. From a theoretical perspective, the *label propagation* algorithm is compelling, as it is based on a network's local properties, which is consistent with sociological notions of communities. In practice, however, modularity optimization procedures, such as *multilevel* or *leading eigenvector*, dominate. Both have been shown to yield results with well-separated communities (Yang et al., 2016), although a resolution limit—the inability of modularity optimization procedures to detect very small communities—has been described (Fortunato and Barthélemy, 2007). Peel et al. (2017) argue that, when choosing a community detection procedure, one should take into consideration both beliefs about the data-generating process and how the outputs will subsequently be used—for example, whether the aim is testing a specific hypothesis or simply coarse-graining the network structure. One may therefore also take the specific features of algorithmic solutions into account. For instance, if we are interested in discovering structural holes, an algorithm, which identifies well-separated communities (i.e., delivers high modularity solutions), is a reasonable choice. If, however, we want to identify communities that enable quick information flow, short paths within the groups are the more important feature.

With the above considerations on data preprocessing and choosing an algorithm, we argue for the primacy of a theory-driven approach; that is, decision-making based on what ties may represent and what researchers are interested in learning. We acknowledge, however, that some decisions will also be data-driven. For example, what constitutes a weak vs. a strong tie will depend on the overall distribution of tie weights and reciprocal ties on the network (Barrat et al., 2004; Serrano et al., 2009). How specific algorithm perform will depend on various properties of a given network, such as its size, its tendency towards

¹ Our study solely focuses on hyperlink ties embedded in the HTML source code of web pages. Other relations, such as textual or visual relations (e.g., logos of affiliated organizations), which are not accompanied by hyperlinks, remain omitted.

² Another possibility that we will not follow up on here is the definition of ordinal threshold categories. That is, instead of retaining tie weights in their original form, researchers may classify weights into groups of weak, medium, and strong ties. This may prevent the illusion of a quasi-metric strength indicator. However, the definition of such categories is, again, in danger of being arbitrary if there is no theoretical reasoning.

Table 1
Capabilities of different algorithms.

Algorithm	Tie weights	Tie direction
Edge betweenness	✓	✓
Multilevel	✓	✗
Leading eigenvector	✓	✗
Label propagation	✓	✗
Map equation	✓	✓

triadic closure, the number of cross-cutting ties between densely connected network modules, the skewness of the degree distribution etc. Excellent guidance for such data-driven criteria is provided by the literature in statistical physics and computer science (Lancichinetti and Fortunato, 2009; Lancichinetti et al., 2008; Yang et al., 2016).

In most cases, neither purely theory-driven nor purely data-driven approaches are likely to be superior. Instead, researchers need to explore the empirical features of their networks with theory in mind; that is, data-driven decision-making must be compatible with the domain-specific notion of community.

Methods

We compare the results of five popular algorithms and of different preprocessing choices using an empirical data set. Our test case is an issue-specific hyperlink network of (mainly) British websites concerned with the topic of food safety, which was collected in July 2014. The network was generated using the *Issue Crawler*³ and eight systematically chosen seed URLs of civil society actors that were deemed central to the issue.⁴ Starting from these websites, a snowball procedure was applied using a crawl depth of two (i.e., internal links were collected from the domain's main site up to two levels) and a degree of separation of one (i.e., from all pages, external links to other websites were followed). For the crawled websites, two levels of internal links were followed, and all ties leading back to websites already in the network were added.

Snowballing is the most inclusive mode of crawling when capturing the interlinking structure of an a priori unknown assemblage of websites (Waldherr et al., 2017). The technique was tested multiple times using varying parameters. While choosing a degree of separation parameter greater than one led to an unmanageable amount of thematically unrelated sites (i.e., many false positives), a crawling depth parameter greater than two yielded insignificantly deviating networks. It is worth noting, that snowball networks do not require an a priori knowledge about their boundaries (Adam et al., 2016). Instead, boundary definition remains an analytical challenge (Maier et al., 2018b). In our case, the boundaries of the network are defined thematically.

As we focused on a network of organizations tied together by hyperlinks rather than on a network of websites, the websites' hosting organizations were manually coded. The network was aggregated so that all websites hosted by the same organization were merged into a single organizational node. Furthermore, the websites were checked for

³ For documentation on Issue Crawler, see: <http://www.issuecrawler.com>

⁴ To define the starting pages of the crawling procedure, we systematically conducted Google searches using different keywords associated with food safety, reviewed the literature, and gathered experts' opinions. This systematic search led to a list of websites, which we checked for availability, update frequency, and centrality of the food safety issue. Finally, we chose the following websites as starting points (seed websites) for the crawl: <http://www.which.co.uk/about-which/what-we-do/which-policy/food/food-safety/> <http://www.consumerfocus.org.uk/wales/policy-research/food> <http://www.sustainweb.org/> <http://www.acornsafety.co.uk/category/food-safety-news-and-advice/> <http://www.greenpeace.org/international/en/campaigns/agriculture/> http://www.cieh.org/policy/food_safety_nutrition.html http://www.foe.co.uk/get-involved/natural_resources.html <http://www.soilassociation.org/>

	Directed	Undirected
Weighted	Directed, weighted network	One undirected, weighted network containing all ties & one network containing only reciprocal ties
Unweighted	/	Unweighted, undirected network with weight thresholds $w = 1, 2, 5$

Fig. 2. Network versions after preprocessing.

thematic relevance using a keyword filtering procedure with *Visual Web Spider*.⁵ The resulting network consisted of 551 organizations and 904 directed and weighted hyperlink ties between them. The network featured a low density of 0.006 and an average degree of 1.76.

Through preprocessing (symmetrizing, dichotomizing), we created the versions of the network represented in Fig. 2. We used *edge betweenness* to assess the impact of varying the treatment of tie weights and directions because it is a flexible algorithm that can handle directed and weighted networks. This allowed us to evaluate the impact of all preprocessing steps while using the same algorithm. After increasing the threshold value for tie weights, we deleted isolates before applying the algorithm. As discussed, it is not possible to settle on an absolute weight threshold, at which ties represent a particular relation type. We chose the tested thresholds so that each step represents major changes in the number of remaining nodes and ties. Increasing the threshold from one to two eliminates those ties most likely to be spurious by focusing on repeated connections. Further, a threshold of five focuses our attention on roughly the top ten percent of weighted ties—the strongest relationships within the network.

To see how different algorithms performed in comparison to the *edge betweenness* algorithm, we applied them to the undirected, weighted network containing all ties. We tested a selection of popular algorithms—namely, the *leading eigenvector*, *multilevel*, and *label propagation* algorithms—on this network. Furthermore, we applied the *map equation* algorithm to the directed, weighted network. We further conducted sensitivity analyses to investigate how varying values for the thresholding technique affects the solutions of the algorithms.⁶

As noted above, the analyses were conducted using the *igraph* package (Csardi and Nepusz, 2006) for the statistical computing environment R (R Development Core Team, 2008). For network visualizations, we used *Gephi* (Bastian et al., 2009).

In order to systematically compare the partitioning solutions, we calculated several measures, including the number of groups, average and median group size, skewness and kurtosis of the distribution of group sizes, the maximum and minimum number of group members, and the number of unassigned nodes and dyads. These measures are descriptive, and there was no objective optimum value. Rather, the aim was to compare how the solutions differ. Furthermore, we used modularity (Newman and Girvan, 2004) and the share of intra-community ties to assess the extent to which communities are separated from each other. For these measures, high values are usually regarded as desirable, as they point toward well-separated communities. We also used

⁵ The keyword filtering was conducted using Visual Web Spider. A detailed documentation of the software may be found here: <http://www.newprosoft.com/web-spider.htm>. A list of the search terms can be found in Table 5 in the Appendix.

⁶ The sensitivity analyses as well as the data set and the respective scripts are available at https://osf.io/sy3r2/?view_only=e133dc22cb8b4511b648aa420fd0deea

the maximum and average diameter to assess reachability between members within communities. Here, low values indicate that all members can be reached via short paths, which is desirable if the theoretical construct under study demands reachability—for example, information flow, user flow, or other flow processes.⁷

Lacking a ground truth variable in our data, we validated results based on a two-step assessment procedure that focused on (a) the *interpretability* of subgroup structures and (b) the *concept fit* of significant subgroups.⁸ In the first step, (a) we asked to what extent the network partitions in fact display significant subgroups. If the aforementioned metrics or other indications suggest that the partitioning does not reflect significant groups, the solution was deemed uninterpretable. Insignificant subgroup structures may be indicated by a high number of communities and a heavy skewness of group size distribution, i.e., there is only one large community and many tiny ones consisting of only one or two nodes. If the solution represents a significant group structure, in the second step, (b) we asked whether the groups correspond to a community concept we are looking for. Two researchers conducted the validity assessment to ensure intersubjectivity. To aid this assessment, we consulted the types of actors and topics within groups,⁹ as well as known institutionalized affiliations (e.g., between different regional chapters of organizations like *Greenpeace* or the *Chartered Institute of Environmental Health*).

Results

We will discuss the results according to the three decisions outlined above. First, we present our findings regarding the different options for handling tie directions. Next, the results regarding tie weights and, finally, the different algorithms are discussed.¹⁰

⁷ For stochastic algorithms, $n = 100$ iterations were run. This pertains to the *label propagation* and *map equation* algorithms. The mean and standard deviation of all measures were calculated.

⁸ Network community detection is a type of unsupervised clustering. Therefore, we used a validity-assessment approach, which was originally developed for interpretations of topic models (Maier et al., 2018a), also an instance of unsupervised clustering.

⁹ The coding of actor type categories was based on a broad categorization similar to that of Rucht et al. (2008) and conducted by two trained coders. The coding reached a satisfactory intercoder reliability of 0.78 (Krippendorff's α). A more detailed description can be found in Miltner et al. (2013). Qualitative inspection was used to gather the main topics that the websites were dealing with.

¹⁰ A file containing all network visualizations with node IDs and a node list is also available at https://osf.io/sy3r2/?view_only=e133dc22cb8b4511b648aa420fd0deea under the file name *Visualizations with IDs and Nodelist* to allow readers to see in more detail which actors are grouped into communities under what conditions.

Table 2
Community solutions depending on tie directionality.

	Undirected network, all ties	Undirected network, reciprocated ties only	Directed network
Number of nodes	551	45	551
Number of ties	904*	45	949*
Number of components	1	5	1
Density	0.006	0.045	0.003
Number of communities	26	7	229
Average number of members	21.19	6.43	2.41
Median number of members	8	2	1
Skewness of community size distribution	2.84	1.36	15.03
Kurtosis of community size distribution	11.05	3.25	226.95
Largest community	158	22	317
Smallest community	2	2	1
Number of single nodes	0	0	224
Number of dyads	3	4	3
Modularity	0.47	0.18	0.13 ^{††}
Share of intra-community ties	0.65	0.91	0.50
Largest diameter	8	3	7
Average diameter	3.69	1.57	2.40 [†]

All values are based on the edge betweenness algorithm (Girvan and Newman, 2002) and were calculated based on weighted networks. “All ties” include all hyperlinks between actors, regardless of direction ($i \rightarrow j, j \rightarrow i, i \leftrightarrow j$). “Reciprocated ties” take only reciprocal hyperlink relations into account ($i \leftrightarrow j$). All values (except for density) were rounded to the second decimal place.

* Differences in the number of ties are due to *igraph* treating reciprocal links as two separate ties in directed networks. As there are 45 reciprocated ties, the directed network contains 45 additional ties.

† When calculating average diameters, values for “communities” consisting of just one member are not taken into account, as they always equal zero and distort the results significantly.

†† Modularity is only implemented for undirected graphs in *igraph*, which must be taken account in the interpretation of these values. However, there is no consensus so far on how to generalize this measure for directed networks (Malliaros and Vazirgiannis, 2013).

Handling tie directionality

The decision on how to handle directions should be the first one. There are three options. The most common one is symmetrizing, which treats all ties (reciprocated and non-reciprocated) equally. The second approach is to settle on reciprocated ties only and the third is to leave the adjacency matrix as is and use the directed network. The results of these three approaches can be found in Table 2. The respective network visualizations are provided in Fig. 3.

The first solution ignores tie directions, thereby creating a network with mostly unreciprocated ties. The community partitions are relatively homogeneously sized (Fig. 3a) and may be interpreted as topically similar groups. For instance, we find groups mainly concerned with contaminated food or with organic food. The solution consists of 26 communities, and the modularity of 0.47 indicates a significant community structure (Newman and Girvan, 2004). The average diameter of 3.69 is rather high, so paths within groups are long in many cases.

The network with only reciprocated ties reveals that there is little reciprocity among the actors’ websites (Fig. 3b). Only 45 nodes maintain reciprocal hyperlinks. The main component contains 37 nodes, which the *edge betweenness* algorithm groups into three communities. The *Chartered Institute of Environmental Health* now forms a community with only its local subdivisions in Ireland and Wales. More generally, the groups exhibit a clear sorting by actor type, with few connections persisting between, for instance, economic and civil society actors. This finding provides evidence that insisting on reciprocity may indeed lead to communities of strategically cooperating actors.

Finally, the solution for the directed network (Fig. 3c) does not yield an interpretable community structure. The algorithm divides the network into 229 communities, 224 of which consist of only one node. There is one large community with 317 nodes. The unbalanced community size distribution is indicated by the large skewness and kurtosis values (Table 2). Although there are large and small groups in every solution, the extreme imbalance indicates a malfunctioning of community detection for the directed network. This is also indicated by the low modularity of 0.13. Not much interpretative value can be derived from this solution: The only sizable group consists of mostly economic

and civil society actors and forms around outlinks from one civil society actor (*Sustain*). While a similar group is found in most solutions, it is much larger and less topically coherent here.

Handling tie weights

Table 3 and Fig. 4 present the community solutions discovered by the *edge betweenness* algorithm for the weighted network, as well as for dichotomous networks with weight threshold values for $w = 1, 2$, and 5. Although the statistical descriptions for the weighted network discussed above (Fig. 4a) and the dichotomous network for $w = 1$ (Fig. 4b) differ, a closer examination of the groups reveals several similarities. The larger, more centralized communities, as well as most pairwise groupings of dyads, are nearly identical in both versions. However, many of the smaller groups in the weighted network are parts of larger communities in the solution for the dichotomized network. Thus, although the number of groups differs, the community assignment is similar. The solution for the dichotomous network may be slightly preferable as it exhibits a lower number of communities and higher modularity. Additionally, the lower skewness and kurtosis of the group size distribution indicate a greater balance among communities.

Regarding the weight threshold models, we observe that even a slight increase of the threshold strongly alters the community structure (Fig. 4c). Taking into account only ties representing at least two hyperlinks shrinks the network to about half its size ($n = 243$ nodes). At $w = 5$, only about one sixth of the nodes and 10.5% of ties remain in the network (Fig. 4d).

With an increased threshold value, community structures show higher modularity values and larger shares of intra-community ties. A closer look reveals consistent patterns of structural changes. While the central actors of large, highly centralized groups remain in their positions, the communities at the network’s periphery change. Some of the small and medium-sized communities fall apart completely or are redistributed across several other groups. Thus, while some groups remain almost the same for $w = 1$ and $w = 2$, we cannot generally assume that increasing the threshold value leads to an equivalent structure in which all groups lose some members equally. Instead, we observe a reorganization of the community structure.

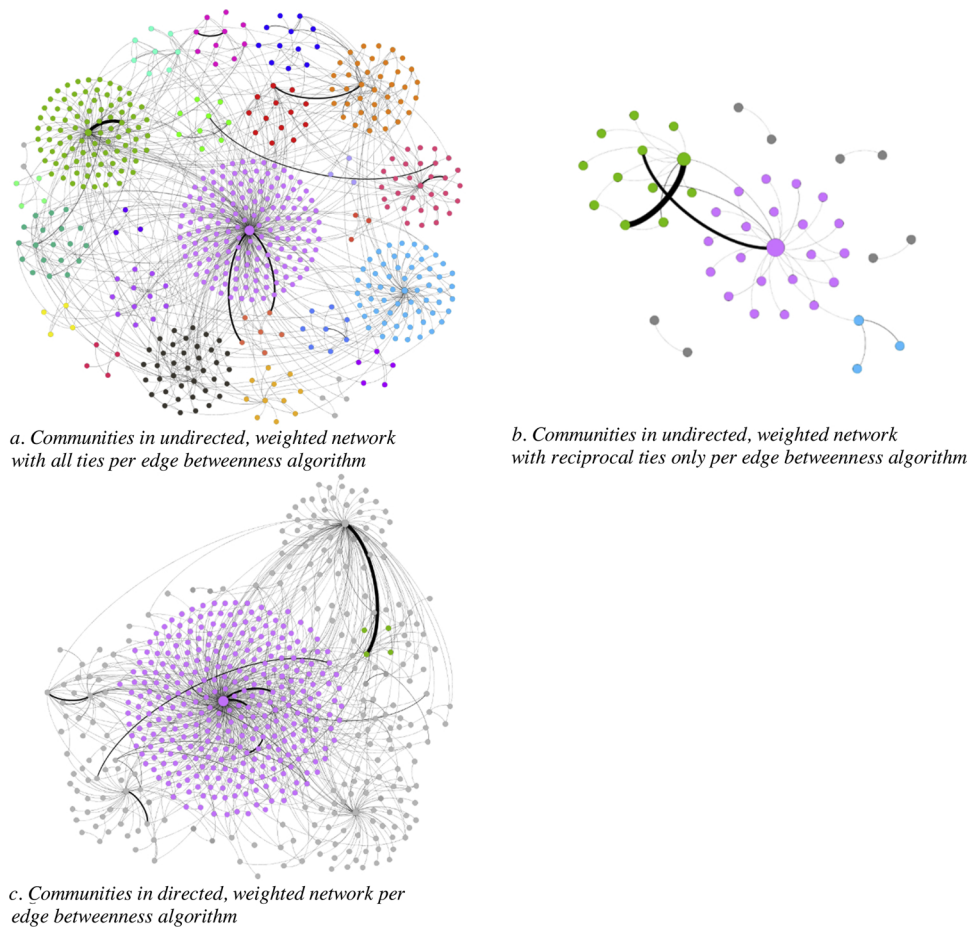


Fig. 3. Community structures for different handling of tie directions.

A more qualitative inspection reveals that an increased weight threshold leads to groups that are sorted more clearly by actor type. For instance, in the weighted network, two central actors (*Sustain* and *Soil Association*) tended to be clustered with many economic actors. This tendency is less pronounced at $w = 2$ and all but disappears at $w = 5$, when both *Sustain* and *Soil Association* form groups with mostly other

civil society actors. Some topical groups (e.g., on organic food) that clearly emerged previously disappear with the increased threshold. Conversely, actors that are marked by a known institutionalized connection (e.g., different regional sections of *Greenpeace*) cluster together more clearly at an increased threshold.

Table 3
Community solutions depending on tie weights.

	Weighted network	Dichotomous network, threshold value $w = 1$	Dichotomous network, threshold value $w = 2$	Dichotomous network, threshold value $w = 5$
Number of nodes	551	551	243	92
Number of ties	904	904	302	95
Number of components	1	1	7	15
Density	0.006	0.006	0.010	0.023
Number of communities	26	18	17	20
Average number of members	21.19	30.61	14.29	4.60
Median number of members	8	12	4	2
Skewness of community size distribution	2.84	2.46	2.22	1.96
Kurtosis of community size distribution	11.05	8.72	7.69	6.08
Largest community	158	182	79	19
Smallest community	2	3	2	2
Number of single nodes	0	0	0	0
Number of dyads	3	0	7	11
Modularity	0.47	0.55	0.64	0.70
Share of intra-community ties	0.65	0.70	0.82	0.84
Largest diameter	8	8	9	3
Average diameter	3.69	4.28	3.29	1.70

All values are based on the edge betweenness algorithm (Girvan and Newman, 2002) and were calculated based on undirected networks. All values (except for density) were rounded to the second decimal place.

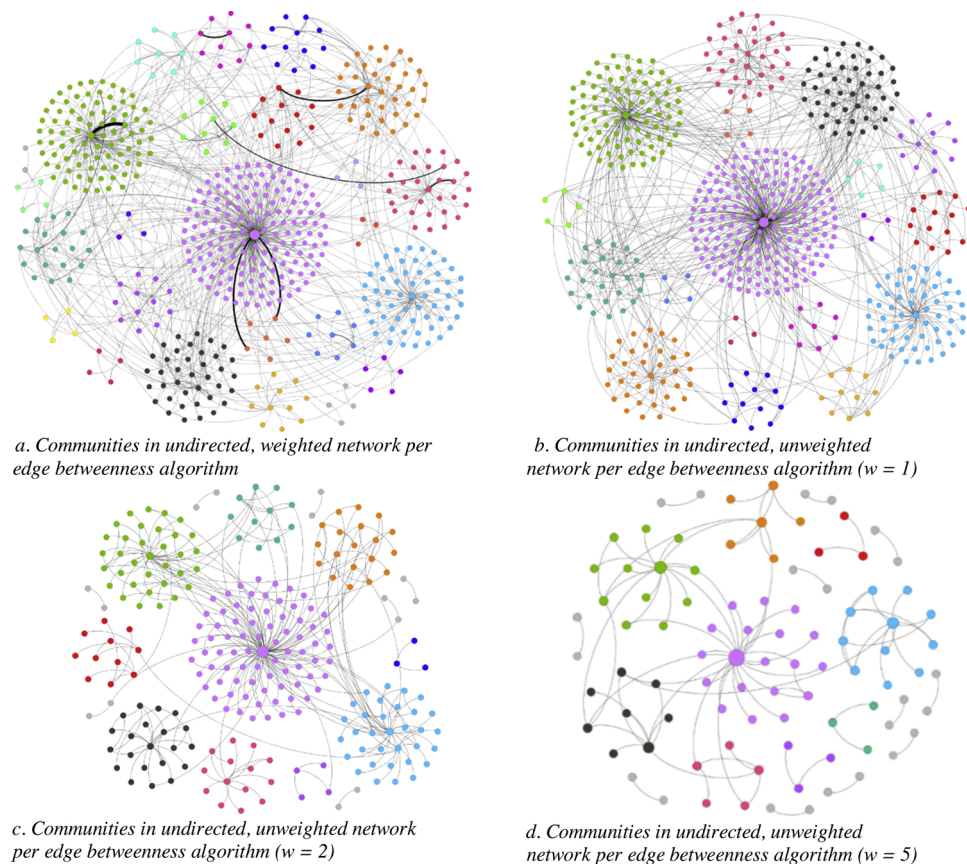


Fig. 4. Community structures for unweighted networks.

Choosing a community detection algorithm

There are numerous community detection algorithms, and it is not obvious how they affect the communities discovered in real-world hyperlink networks. We compared the results of several algorithms against the *edge betweenness* algorithm (applied above)—namely, *map equation* for directed networks, as well as *leading eigenvector*, *multilevel*, and *label propagation* for undirected networks. Comparative results are provided in Table 4 and Fig. 5.

Focusing on the right side of Table 4 (directed networks), we observe that the *map equation* algorithm, like *edge betweenness*, reveals problems when searching for communities in the directed network. As the underlying procedure is based on the movements of a random walker, communities should be detected only in the areas, in which the walker lingers. At first glance, this algorithm appears to find an interpretable community solution. As expected for a directed network, the algorithm leaves some nodes and dyads unassigned. However, on average 28.89 communities with at least three and up to 288.83 nodes remain, and the mean modularity of 0.56 suggests a significant community structure.

However, there are some inconsistencies. The algorithm tends to place central nodes outside densely connected network areas leading to communities that are internally disconnected. This is inconsistent with conceptual notions of communities in networks. Elsewhere in the network, the algorithm shows results consistent with these criteria, removing actors with no outlinks from groups. This is desirable, as these actors do not allow for further movement within the community in the case of directed networks. Qualitatively, the groups tend to be less topically coherent, with, for instance, the previously identified group on food contamination merging with the largest group in the network and the organic food actors being distributed across several communities. Therefore, the directed network does not seem to be more capable of

finding topically coherent groups.

All algorithms applied to the *undirected network* display significant community structures. However, the solutions also exhibit substantial differences, proving how critical the choice of an algorithm is. Although *edge betweenness* is very popular, its solution is the one with the lowest modularity score and highest average and absolute diameter. The number of communities is comparable to that of *leading eigenvector* and *label propagation*. The *multilevel* algorithm's solution combines the lowest number of groups with the highest modularity score and features the most balanced community size distribution.

A more detailed examination reveals that *leading eigenvector* divides the network into communities with remarkably low diameters. Each node can reach every other node in a community within a maximum path length of four, the average being only 2.17. This is interesting because short diameters are not an optimization criterion of the algorithm. However, as with the *map equation* algorithm, this procedure yields an internally unconnected group in one case. Again, this is inconsistent with notions of what constitutes a group in a network. Regarding modularity, *leading eigenvector* (0.65) performs slightly worse compared to *label propagation* (0.70) and *multilevel* (0.73). Qualitatively, many findings hold in comparison to *edge betweenness*. Topical groups are still visible. Known institutionalized connections (e.g., different *Greenpeace* or *Chartered Institute* chapters) are reflected well in the group structure.

The *multilevel* algorithm performs well according to several criteria. It has the highest modularity value (0.73) and lowest number of groups (17). Regarding average diameters, it performs worse (3.29) than *leading eigenvector* (2.17) and *label propagation* (2.23). These high diameters point to long path distances between group members; however, they are at least partly a result of the larger community sizes. Focusing on individual groups (Fig. 5b) reveals a tendency to merge small groups discovered by other algorithms into larger ones. This is likely an effect

Table 4
Community solutions of different algorithms.

	Edge betweenness (Girvan and Newman, 2002), undirected	Leading eigenvector (Newman, 2006b), undirected	Multilevel (Blondel et al., 2008), undirected	Label propagation** (Raghavan et al., 2007), undirected	Edge betweenness (Girvan and Newman, 2002), directed	Map equation** (Rosvall et al., 2009), directed
Number of nodes	551	551	551	551	551	551
Number of ties	904*	904†	904†	904*	949‡	949‡
Number of components	1	1	1	1	1	1
Density	0.006	0.006	0.006	0.006	0.003	0.003
Number of communities	26	29	17	29.01 (1.45)	229	73.76 (13.71)
Average number of members	21.19	19	32.41	19.03 (0.93)	2.41	7.78 (1.70)
Median number of members	8	3	9	3.07 (0.20)	1	1.91 (0.63)
Skewness of community size distribution	2.84	3.88	2.60	4.14 (0.16)	15.03	8.08 (0.79)
Kurtosis of community size distribution	11.05	18.16	9.13	20.09 (1.34)	226.95	68.45 (12.74)
Largest community	158	235	232	258.21 (4.85)	317	288.83 (21.44)
Smallest community	2	1	2	1.89 (0.31)	1	1.01 (0.10)
Number of single nodes	0	5	0	0.12 (0.36)	224	31.89 (13.22)
Number of dyads	3	8	3	8.1 (0.88)	3	12.98 (0.47)
Modularity	0.47	0.65	0.73	0.70 (0.01)	0.13††	0.56 (0.02)††
Share of intra-community ties	0.65	0.65	0.74	0.71 (0.01)	0.50	0.42 (0.10)
Largest diameter	8	4‡	7	5.73 (0.72)	7	7.41 (0.65)‡
Average diameter	3.69	2.17†,‡	3.29	2.31 (0.08)	2.40†	2.28 (0.06)†,‡

All values were calculated based on weighted networks. The first four columns are based on undirected networks, and the last two are based on directed networks. All values (except for density) were rounded to the second decimal place.

* Differences in the number of ties are due to *igraph* treating reciprocal links as two separate ties in directed networks. As there are 45 reciprocated ties, the directed network contains 45 additional ties.

** The algorithm contains stochastic elements. The presented values refer to the mean of 100 runs. The respective standard deviations are indicated in brackets.

† When calculating average diameters, values for “communities” consisting of just one member are not taken into account, as they always equal zero and distort the results significantly.

†† Modularity is only implemented for undirected graphs in *igraph*, which must be taken account in the interpretation of these values. However, there is no consensus so far on how to generalize this measure for directed networks (Malliaros and Vazirgiannis, 2013).

‡ These algorithms discovered groups that were not internally connected. In these cases, the diameters for individual components were taken into account.

of a resolution limit that prohibits the detection of very small groups (Fortunato and Barthélemy, 2007). As the clusters are larger, they are also less easily interpretable in terms of shared topics or identities. However, actors among whom there is a known institutional affiliation are again clearly grouped together. Moreover, we find some clusters dominated by a single actor type. Therefore, we have indication that the grouping is meaningful, but would require more sophisticated procedures to decode regarding shared topics.

The *label propagation* algorithm achieves a high average modularity of 0.70. What is more, the largest (5.73) and mean average diameter (2.32) are also comparatively small. On average, the algorithm divides the network into 29.01 communities, 8.1 of which consist of only 2 members. Among the remaining groups, many contain only few nodes. These small communities are usually grouped together by other algorithms too, but they are merged with larger communities. However, the *label propagation* algorithm also finds the largest community detected by any algorithm in the undirected network, revealing a highly skewed community size distribution. The small communities allow additional insights into topical alignments, as we find, for instance, one community exclusively concerned with organic baking.

In additional analyses, we also checked how varying global thresholding values affected the partitioning solutions of the algorithms (see Footnote 6). Generally, with higher global thresholds for tie weights, we witness an increase in modularity and in the share of intra-community ties. Conversely, the average diameter and the average number of community members decrease with an increasing threshold. The number of communities remains relatively stable for all algorithms except one. For the *edge betweenness* algorithm (applied on the directed network), the number of communities decreases in an exponential-like trajectory.

Discussion

As we have seen, communities, which are vaguely defined as densely connected parts of a network, can be specified in a multitude of ways. We discussed three general decisions that researchers have to make concerning community detection: How do we handle (1) the directionality and (2) the weight of ties? And (3) which community detection procedure do we apply? We will first discuss the implications of our empirical findings before connecting them to our theoretical considerations.

Empirical consequences of different community detection procedures

The standard case in current research practice is the undirected, weighted network, which contains both reciprocal and non-reciprocal ties. With our empirical example, we were able to specify the consequences of disregarding tie directions. Keeping only reciprocal ties drastically reduced the size of our network. Barely one tenth of the actors remained in the resulting network, and the community structure had little in common with the one discovered in a non-reciprocal network. Interpreting non-reciprocal ties as proof of reciprocal social relations (e.g., strategic alliances) between actors in a community may therefore overestimate the degree of cooperation. Conversely, the reciprocal network offers a conservative perspective on the degree of interlinkage and points out only highly connected actor groups. This comes at the cost of being unable to draw conclusions about most actors from the original network.

Both the *edge betweenness* and *map equation* algorithm revealed issues when applied to a directed hyperlink network. The former found one large group, supplemented by a large number of unassigned nodes, while the latter exhibited a tendency to place central nodes outside of

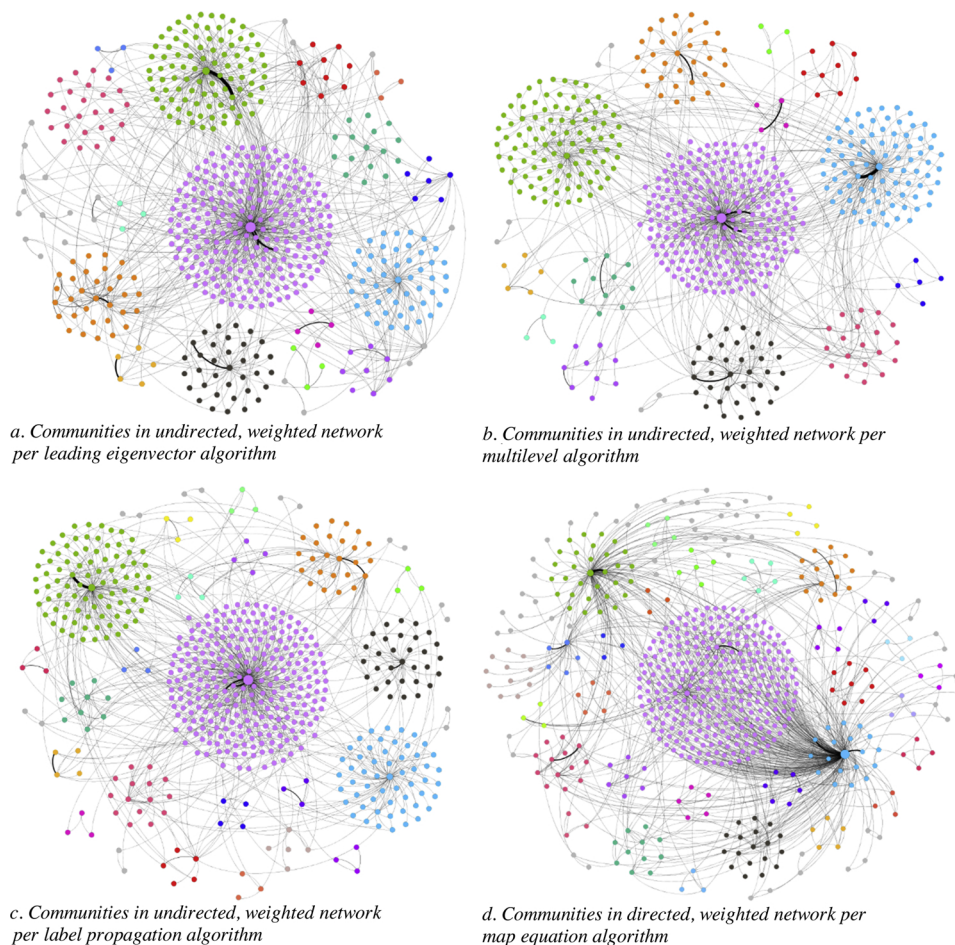


Fig. 5. Community structure discovered by different algorithms.

communities. As cohesive groups within networks are conceptualized as constellations of actors densely connected by ties (Wasserman and Faust, 2009), such a community partitioning challenges the very basis of that definition.

Altogether, the findings regarding the directed network may signify either a weakness of the algorithms or an absence of internally well-connected communities in our network. Only replications on different networks can provide a conclusive answer to this conundrum, but the latter may well be the case. Like most hyperlink networks, ours is sparse and has few reciprocal ties. A small group of actors sends and receives most of the hyperlinks, leading to star-configurations. These features may prohibit lingering in network areas, and this becomes evident only when the direction of hyperlinks is considered. In any event, this finding indicates that researchers ought to be very cautious about making assumptions about directed processes, such as user traffic, based on undirected networks.

Concerning tie weights, communities were similar regardless of whether we dichotomized the network before running a community detection procedure. Increasing the weight threshold even slightly, however, led to entirely different communities in parts of the network. This is a warning sign against overinterpreting communities detected in the weighted network, as many are based on very weak hyperlink relations. Such communities may not be appropriate indicators for social phenomena, such as shared ideology or strategic cooperation. If one is merely interested in topical similarity, however, even infrequent links may be meaningful.

Different algorithms applied to the same network yielded quite different community solutions, with strengths and weaknesses emerging. The *edge betweenness* algorithm is flexible regarding tie weights

and directions and delivered a solution with many medium-sized communities; however, compared to other algorithms, it performed poorly regarding modularity and diameter lengths. *Multilevel* found a solution with a small number of groups and a high modularity, but also high diameters. *Leading eigenvector* performed best regarding diameter lengths, but left many nodes unassigned and, in one case, discovered an internally unconnected community. *Label propagation* combined low diameters and a high modularity, identifying a large number of very small groups in the process.

In summary, while large, highly centralized communities could be detected consistently, divergent results were found for peripheral groups, which were small and lacked central hub nodes. We, therefore, advocate for a theory-based approach to community detection, as well as for better reporting standards to explicitly justify methodological choices.

Theory-based decisions on community detection methods

Why theory matters

In contrast to studies working with networks where a community ground truth is available, our data set incorporates natural hurdles that researchers face when exploring community structures. Notwithstanding the fact that “no algorithm can uniquely solve community detection (...) and that there can be no algorithm that is optimal for all possible community detection tasks” (Peel et al., 2017, p. 1), defining an *ever-valid* ground truth independently from the substantive theoretical notion of a community at hand is in our view largely impossible. In our case, a ‘true’ community structure could be defined only for strategically cooperating actors where publicly available documents

exist that state the existence of their alliance. Informal collectives, such as movements, often lack documents stating their cooperation formally. In larger discourse coalitions, actors might not be aware of everyone with whom they share discursive positions. Ground truth changes with the theoretical perspective.

Moreover, we face a different ontological problem when turning to discursive coalitions, which *emerge* from their communicative relations. This ontological problem turns the ground truth challenge upside down: Instead of asking if the solution reflects true communities there are no true communities in the first place. Thus, we strongly require theoretical guidance to define when a condensed communicative structure signifies a community.

A theoretical guide

Based on our review of studies broadly concerned with civil society, we found four prevailing theoretical interpretations of communities in hyperlink networks. These interpretations demand different levels of cohesion for communities. Fig. 6 sums up our recommendations for choosing a community detection procedure. The aim of providing such a heuristic scheme is to provide a concise overview of our main findings. It does not replace thinking critically about how to best approximate any particular theoretical group concept.

First, researchers need to explicate what theoretical concept they believe the hyperlinks in a network signify and whether it is a strong or weak type of relation. Topical similarity, ideological association, and strategic alliances (i.e., our three diagnostic concepts) represent progressively stronger types of relationships, which ought to be reflected in the level of cohesion of subgroups. Furthermore, prognostically, hyperlinks may signify potential avenues of user flow.

Next, the way of handling tie directions must be considered. If the theoretical concept is directed (such as potential user traffic), tie directions ought to be retained, even taking into account the possible issues of community detection in directed networks. If the theoretical concept is reciprocal, we recommend limiting the analysis to reciprocal ties, even if it shrinks the network. This is particularly vital for strategic alliances or advocacy coalitions, but even with ideological associations or discourse coalitions, reciprocal relations may present more compelling evidence. Rather than treating the undirected non-reciprocal network as the default, it should only be used if it is in line with the construct under study (e.g., topical similarity).

For tie weights, too, theoretical considerations should be paramount. For theoretical concepts representing strong relations, such as ideological associations or strategic alliances, thresholding makes sense to foreground frequent connections and avoid overinterpreting spurious links. If, however, more frequent linkages represent stronger relations, using a weighted network makes sense. User traffic is an illustrative

example, as users are more likely to click through from one website to another if there are more links. While these are theory-driven approaches, we acknowledge that they must be combined with data-driven considerations. For example, if no connections remain in the network at a certain threshold, researchers may need to settle for a lower one. However, if all connections in the network are weak, one should be cautious about drawing conclusions on strong relational concepts.

Regarding algorithm choice, more deliberation is necessary. For directed networks, researchers are limited to the *map equation* and *edge betweenness* algorithms in *igraph* (Csardi and Nepusz, 2006). For undirected networks, a bigger selection of procedures is available. Here, researchers may focus on two aspects: how the algorithm’s principles relate to the theoretical concept and what the aim of data analysis is. One main difference can be drawn between agglomerative (*label propagation, multilevel*) and divisive (*edge betweenness, leading eigenvector*) algorithms. One may argue that agglomerative algorithms more closely mirror processes of network generation and therefore the mechanisms behind actor relations (diagnostic interpretations).

Regarding the aim of data analysis, we may focus on the empirical features the algorithms produce. If, for instance, the aim is to identify structural holes, an algorithm that produces a high modularity is a reasonable choice. If researchers want to examine communication processes, short group diameters should enable such flows. Granularity may also be considered; that is, searching for cooperating actor groups may work better with an algorithm without a resolution limit, one that can identify small, tightly connected actor groups. Finally, there is value in trying more than one approach to community detection, as this reveals which findings are robust against methodological variations.

Can we generalize?

We chose real-world observational data of civil society online communication about food safety. Therefore, our empirical test case is customized for studying hyperlink relations among civil society actors and other important social actor types. Any generalizations beyond this domain have to be carefully considered and adapted to the specifics of the networks and theoretical background.

For instance, community detection in social media networks might be tentatively informed by our scheme. Friendship or follower networks also represent an infrastructure for information flow granting visibility and endorsement to highly followed individuals. Similar mechanisms such as homophily, triadic closure, and resource dependence lead to communicative connectivity among actors and thus to the emergence of community structures (Guilbeault et al., 2018; Ugander et al., 2012). However, most nodes of such networks are individuals with a less stable and strategic agenda than corporate actors in hyperlink networks

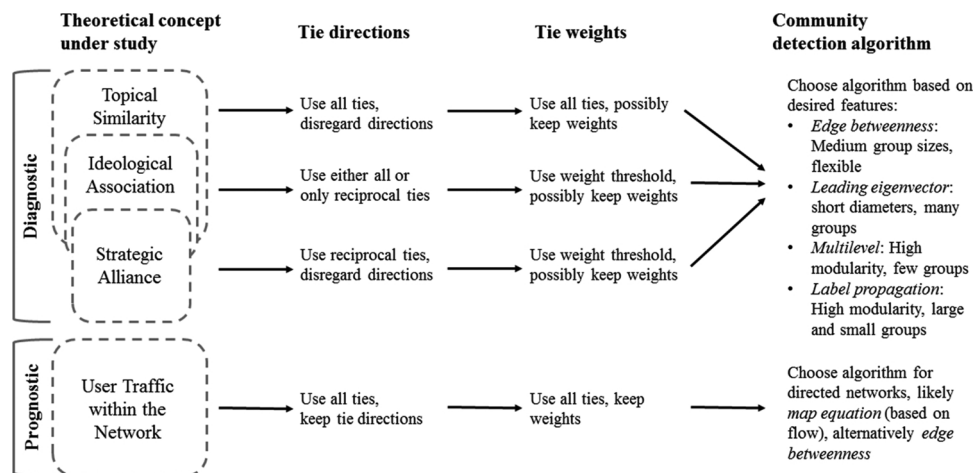


Fig. 6. Theory-based decisions in community detection in hyperlink networks.

(Lotan et al., 2011). Consequently, social media network relations may be more fluid and less enduring than hyperlink connections.

Even more caution is advised when transferring our conclusions to networks of offline social relationships, such as friendship networks among students or networks of war among nation states. Other theories apply to such networks and, in accordance with them, other algorithms and preprocessing steps may be useful.

This paper offers an example of how a scheme of the meaning of ties can be abstracted from a review of domain-specific literature and connected to concrete methodological choices. Similar schemes may be derived from the literature in other research fields, be they online (e.g., social media influencers, bloggers, news media) or offline networks (e.g., protests, friendships, trade). The general message to take away from our paper is that the most appropriate partitioning of a social network is guided by a substantial social-scientific concept.

Directions for future research

Some questions remain. Within the scope of this paper, we could not test every combination of the three variable conditions outlined. Further, the qualitative assessment of communities' interpretability could not be discussed in full breadth here. A more systematic approach (e.g., by combining community detection and topic modeling) could be fruitful.

Additionally, there are further possibilities of community detection that we could not take into account. First, there are other algorithms with the same aims as the ones tested. Second, there is the question of allowing for overlap between communities, enabling researchers to identify actors who belong to more than one group. With any case study, one must be cautious to generalize findings. Similar work on different networks should be fruitful in grounding or qualifying our findings.

Funding

This publication was created in the context of the Research Unit "Political Communication in the Online World" (1381), subproject 7, which was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation). The subproject was also funded by the Swiss National Science Foundation (SNF). Work on this publication was also partly funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project number 290045248 – SFB 1265.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.socnet.2019.07.001>.

References

Ackland, R., O'Neil, M., 2011. Online collective identity: the case of the environmental movement. *Soc. Networks* 33 (3), 177–190. <https://doi.org/10.1016/j.socnet.2011.03.001>.

Adam, S., Häußler, T., Schmid-Petri, H., Reber, U., 2016. Identifying and analyzing hyperlink issue networks. In: Vowe, G., Henn, P. (Eds.), *Political Communication in the Online World. Theoretical Approaches and Research Designs*. Routledge, New York, pp. 233–247.

Adam, S., Häußler, T., Schmid-Petri, H., Reber, U., 2018. Political contestation online: analyzing coalitions and their online strength in the field of climate change. May. Paper Presented at the 68th Annual Conference of the International Communication Association (ICA).

Adamic, L.A., 1999. The small world web. In: Abiteboul, S., Vercoustre, A.M. (Eds.), *Research and Advanced Technology for Digital Libraries*. ECDL 1999. Lecture Notes in Computer Science, vol. 1696 Springer VS, Berlin.

Adamic, L.A., Glance, N., 2005. The political blogosphere and the 2004 US election: divided they blog. August. Proceedings of the 3rd International Workshop on Link Discovery 36–43. <https://doi.org/10.1145/1134271.1134277>.

Baldassarri, D., Diani, M., 2007. The integrative power of civic networks. *Am. J. Sociol.* 113 (3), 735–780. <https://doi.org/10.1086/521839>.

Barabási, A.-L., 2016. *Network Science*. Cambridge University Press, Cambridge.

Barrat, A., Barthélemy, M., Pastor-Satorras, R., Vespignani, A., 2004. The architecture of complex weighted networks. *Proc. Natl. Acad. Sci.* 101 (11), 3747–3752.

Barzilai-Nahon, K., 2008. Toward a theory of network gatekeeping: a framework for exploring information control. *J. Am. Soc. Inf. Technol.* 59 (9), 1493–1512. <https://doi.org/10.1002/asi.20857>.

Bastian, M., Heymann, S., Jacomy, M., 2009. Gephi: an open source software for exploring and manipulating networks. *Proceedings of the Third International ICWSM Conference*. pp. 361–362.

Benkler, Y., Roberts, H., Faris, R., Solow-Niederman, A., Etling, B., 2015. Social mobilization and the networked public sphere: mapping the SOPA-PIPA debate. *Polit. Commun.* 32 (4), 594–624. <https://doi.org/10.1080/10584609.2014.986349>.

Bennett, W.L., Foot, K., Xenos, M., 2011. Narratives and network organization: a comparison of fair trade systems in two nations. *J. Commun.* 61 (2), 219–245. <https://doi.org/10.1111/j.1460-2466.2011.01538.x>.

Bennett, W.L., Segerberg, A., 2013. *The Logic of Connective Action: Digital Media and the Personalization of Contentious Politics*. Cambridge University Press, New York.

Blondel, V.D., Guillaume, J.-L., Lambiotte, R., Lefebvre, E., 2008. Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* 2008 (10), 1–12. <https://doi.org/10.1088/1742-5468/2008/10/P10008>.

Borgatti, S.P., Everett, M.G., Johnson, J.C., 2013. *Analyzing Social Networks*. Sage, London.

Brin, S., Page, L., 1998. The anatomy of a large-scale hypertextual web search engine. *Comput. Netw. Insn Syst.* 30 (1–7), 107–117. [https://doi.org/10.1016/S0169-7552\(98\)00110-X](https://doi.org/10.1016/S0169-7552(98)00110-X).

Bruns, A., 2007. Methodologies for mapping the political blogosphere: an exploration using the IssueCrawler research tool. *First Monday* 12 (5). <https://doi.org/10.5210/fm.v12i5.1834>.

Burris, V., Smith, E., Strahm, A., 2000. White supremacist networks on the Internet. *Sociol. Focus* 33 (2), 215–235. <https://doi.org/10.1080/00380237.2000.10571166>.

Castells, M., 2010. *The Information Age: Economy, Society, and Culture. Volume I: the Rise of the Network Society*, 2nd ed. Blackwell Publishing, Chichester.

Centola, D., 2013. Homophily, networks, and critical mass: solving the start-up problem in large group collective action. *Ration. Soc.* 25 (1), 3–40. <https://doi.org/10.1177/1043463112473734>.

Csardi, G., Nepusz, T., 2006. The igraph software package for complex network research. *InterJournal Complex Systems* Retrieved from. <http://igraph.org>.

De Maeyer, J., 2013. Towards a hyperlinked society: a critical review of link studies. *New Media Soc.* 15 (5), 737–751. doi:0.1177/1461444812462851.

Diani, M., 1992. The concept of social movement. *Sociol. Rev.* 40 (1), 1–25. <https://doi.org/10.1111/j.1467-954X.1992.tb02943.x>.

Fortunato, S., 2010. Community detection in graphs. *Phys. Rep.* 486 (3–5), 75–174. <https://doi.org/10.1016/j.physrep.2009.11.002>.

Fortunato, S., Barthélemy, M., 2007. Resolution limit in community detection. *Proc. Natl. Acad. Sci. U. S. A.* 104 (1), 36–41. <https://doi.org/10.1073/pnas.0605965104>.

Ghasemian, A., Hosseinmardi, H., Clauset, A., 2018. Evaluating Overfit and Underfit in Models of Network Community Structure. Preprint arXiv:1802.10582v2.

Girvan, M., Newman, M.E.J., 2002. Community structure in social and biological networks. *Proc. Natl. Acad. Sci.* 99 (12), 7821–7826. <https://doi.org/10.1073/pnas.122653799>.

González-Bailón, S., 2009. Opening the black box of link formation: social factors underlying the structure of the web. *Soc. Networks* 31 (4), 271–280. <https://doi.org/10.1016/j.socnet.2009.07.003>.

González-Bailón, S., Wang, N., 2016. Networked discontent: the anatomy of protest campaigns in social media. *Soc. Networks* 44, 95–104. <https://doi.org/10.1016/j.socnet.2015.07.003>.

Guilbeault, D., Becker, J., Centola, D., 2018. Complex contagions: a decade in review. In: Lehmann, S., Ahn, Y. (Eds.), *Spreading Dynamics in Social Systems*. Springer, Cham, pp. 3–25. https://doi.org/10.1007/978-3-319-77332-2_1.

Hajer, M.A., 1995. *The Politics of Environmental Discourse: Ecological Modernization and the Policy Process*. Oxford University Press, Oxford, UK.

Hartman, R., Faustino, J., Pinheiro, D., Menezes, R., 2017. Assessing the suitability of network community detection to available meta-data using rank stability. August. Proceedings of the International Conference on Web Intelligence 162–169. <https://doi.org/10.1145/3106426.3106493>.

Herring, S.C., Kouper, I., Paolillo, J.C., Scheidt, L.A., Tyworth, M., Welsch, P., Wright, E., Yu, N., 2005. Conversations in the blogosphere: an analysis "from the bottom up." Proceedings of the 38th Hawaii International Conference on System Sciences 107b. <https://doi.org/10.1109/HICSS.2005.167>.

Hyun, K.D., 2012. Americanization of web-based political communication? A comparative analysis of political blogospheres in the United States, the United Kingdom, and Germany. *Journal. Mass Commun.* Q. 89 (3), 397–413. <https://doi.org/10.1177/1077699012447919>.

Kaiser, J., Puschmann, C., 2017. Alliance of antagonism: counterpublics and polarization in online climate change communication. *Commun. Public* 2 (4), 371–387. <https://doi.org/10.1177/2057047317732350>.

Kleinberg, J., Lawrence, S., 2001. The structure of the web. *Science* 294 (5548), 1849–1850. <https://doi.org/10.1126/science.1067014>.

Kriesi, H., Koopmans, R., Duyvendak, J.W., Giugni, M.G., 1992. New social movements and political opportunities in Western Europe. *Eur. J. Polit. Res.* 22, 219–244. <https://doi.org/10.1111/j.1475-6765.1992.tb00312.x>.

Lancichinetti, A., Fortunato, S., 2009. Community detection algorithms: a comparative analysis. *Phys. Rev. E* 80 (5), 056117.

Lancichinetti, A., Fortunato, S., Radicchi, F., 2008. Benchmark graphs for testing community detection algorithms. *Phys. Rev. E* 78 (4), 046110.

Leskovec, J., Lang, K.J., Mahoney, M., 2010. Empirical comparison of algorithms for network community detection. April. Proceedings of the 19th International

- Conference on World Wide Web 631–640. <https://doi.org/10.1145/1772690.1772755>.
- Lotan, G., Graeff, E., Ananny, M., Gaffney, D., Pearce, I., boyd, d., 2011. The revolutions were tweeted: information flows during the 2011 Tunisian and Egyptian revolutions. *Int. J. Commun. 5*, 1375–1405.
- Maier, D., Waldherr, A., Miltner, P., Wiedemann, G., Niekler, A., Keinert, A., Pfetsch, B., Heyer, G., Reber, U., Häussler, T., Schmid-Petri, H., Adam, S., 2018a. Applying LDA topic modeling in communication research: Toward a valid and reliable methodology. *Commun. Methods Meas. 12* (2–3), 93–118. <https://doi.org/10.1080/19312458.2018.1430754>.
- Maier, D., Waldherr, A., Miltner, P., Jähnichen, P., Pfetsch, B., 2018b. Exploring issues in a networked public sphere: Combining hyperlink network analysis and topic modeling. *Social Science Computer Review 36* (1), 3–20. <https://doi.org/10.1177/0894439317690337>.
- Malliaros, F.D., Vazirgiannis, M., 2013. Clustering and community detection in directed networks: a survey. *Phys. Rep. 533* (4), 95–142. <https://doi.org/10.1016/j.physrep.2013.08.002>.
- McCarthy, J.D., Zald, M.N., 1977. Resource mobilization and social movements: a partial theory. *Am. J. Sociol. 82* (6), 1212–1241.
- Miltner, P., Maier, D., Pfetsch, B., Waldherr, A., 2013. Onlinenetworks of civil society actors as an indicator for politicization? A hyperlink analysis of the food safety issue in Germany. *Catalan Journal of Communication & Cultural Studies 5* (2), 201–220. https://doi.org/10.1386/cjcs.5.2.201_1.
- Monge, P.R., Contractor, N.S., 2003. *Theories of Communication Networks*. Oxford University Press, Oxford.
- Newman, M.E.J., 2006. Finding community structure in networks using the eigenvectors of matrices. *Phys. Rev. E 74* (3), 1–22. <https://doi.org/10.1103/PhysRevE.74.036104>.
- Newman, M.E.J., Girvan, M., 2004. Finding and evaluating community structure in networks. *Phys. Rev. E 69* (2), 1–15. <https://doi.org/10.1103/PhysRevE.69.026113>.
- Nitschke, P., Donges, P., Schade, H., 2014. Political organizations' use of websites and Facebook. *New Media Soc. 18* (5), 744–764. <https://doi.org/10.1177/1461444814546451>.
- Papadopoulos, S., Kompatsiaris, Y., Vakali, A., Spyridonos, P., 2012. Community detection in social media. *Data Min. Knowl. Discov. 24* (3), 515–554. <https://doi.org/10.1007/s10618-011-0224-z>.
- Park, H.W., 2003. Hyperlink network analysis: a new method for the study of social structure on the web. *Connections 25* (1), 49–61.
- Park, H.W., Thelwall, M., 2003. Hyperlink analyses of the world wide web: a review. *J. Comput. Commun. 8* (4). <https://doi.org/10.1111/j.1083-6101.2003.tb00223.x>.
- Park, H.W., Thelwall, M., 2008. Link analysis: hyperlink patterns and social structure on politicians' websites in South Korea. *Qual. Quant. 42* (5), 687–697. <https://doi.org/10.1007/s11135-007-9109-z>.
- Peel, L., Larremore, D.B., Clauset, A., 2017. The ground truth about metadata and community detection in networks. *Sci. Adv. 3* (5), e1602548. <https://doi.org/10.1126/sciadv.1602548>.
- Pfetsch, B., Adam, S., Bennett, W.L., 2013. The critical linkage between online and offline media. *Javn. - Public 20* (3), 9–22. <https://doi.org/10.1080/13183222.2013.11009118>.
- Pfetsch, B., Maier, D., Miltner, P., Waldherr, A., 2016. Challenger networks of food policy on the Internet: a comparative study of structures and coalitions in Germany, the UK, the US, and Switzerland. *Int. J. of E-Politics (IJEP) 7* (1), 16–36. <https://doi.org/10.4018/IJEP.2016010102>.
- Pilny, A., Shumate, M., 2012. Hyperlinks as extensions of offline instrumental collective action. *Inf. Commun. Soc. 15* (2), 260–286. <https://doi.org/10.1080/1369118X.2011.606328>.
- R Development Core Team, 2008. *R: a Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna.
- Raghavan, U.N., Albert, R., Kumara, S., 2007. Near linear time algorithm to detect community structures in large-scale networks. *Phys. Rev. E 76* (3), 036106. <https://doi.org/10.1103/PhysRevE.76.036106>.
- Rogers, R., Marres, N., 2000. Landscaping climate change: a mapping technique for understanding science and technology debates on the world wide web. *Public Underst. Sci. 9* (2), 141–163. <https://doi.org/10.1088/0963-6625/9/2/304>.
- Rosvall, M., Axelsson, D., Bergstrom, T.C., 2009. The map equation. *Eur. Phys. J. Spec. Top. 178* (1), 13–23. <https://doi.org/10.1140/epjst/e2010-01179-1>.
- Rucht, D., Yang, M., Zimmermann, A., 2008. *Politische Diskurse Im Internet Und in Zeitungen: Das Beispiel Genfood. [Political Discourse on the Internet and in Newspapers: the Example of Genetically Modified Food.]* Wiesbaden: VS.
- Sabatier, P.A., 1988. An advocacy coalition framework of policy change and the role of policy-oriented learning therein. *Policy Sci. 21* (2–3), 129–168. <https://doi.org/10.1007/BF00136406>.
- Serrano, M.Á., Boguná, M., Vespignani, A., 2009. Extracting the multiscale backbone of complex weighted networks. *Proc. Natl. Acad. Sci. 106* (16), 6483–6488.
- Sereno, K., 2010. *Understanding the Hyperlinks Politics Better: the Israeli Protest Networks As a Case Study. Paper Presented at Internet, Politics, Policy 2010. September. An Impact Assessment*, Oxford.
- Shumate, M., Dewitt, L., 2008. The north/south divide in NGO hyperlink networks. *J. Comput. Commun. 13* (2), 405–428. <https://doi.org/10.1111/j.1083-6101.2008.00402.x>.
- Shumate, M., Lipp, J., 2008. Connective collective action online: an examination of the hyperlink network structure of an NGO issue network. *J. Comput. Commun. 14* (1), 178–201. <https://doi.org/10.1111/j.1083-6101.2008.01436.x>.
- Snow, D.A., Rochford, E.B.J., Worden, S.K., Benford, R.D., 1986. Frame alignment processes, micromobilization, and movement participation. *Am. Sociol. Rev. 51* (4), 464–481. <https://doi.org/10.2307/2095581>.
- Ugander, J., Backstrom, L., Marlow, C., Kleinberg, J., 2012. Structural diversity in social contagion. *PNAS 109* (16), 5962–5966. <https://doi.org/10.1073/pnas.1116502109>.
- Waldherr, A., Maier, D., Miltner, P., Günther, E., 2017. Big data, bignoise: The challenge of finding issue networks on the web. *Social Science Computer Review 35* (4), 427–443. <https://doi.org/10.1177/0894439316643050>.
- Wasserman, S., Faust, K., 2009. *Social Network Analysis: Methods and Applications*, 18th ed. Cambridge University Press, Cambridge.
- Wu, L., Ackland, R., 2014. How web 1.0 fails: the mismatch between hyperlinks and clickstreams. *Soc. Netw. Anal. Min. 4* (1), 1–7. <https://doi.org/10.1007/s13278-014-0202-8>.
- Yang, Z., Algesheimer, R., Tessone, C.J., 2016. A comparative analysis of community detection algorithms on artificial networks. *Sci. Rep. 6*, 1–16. <https://doi.org/10.1038/srep30750>.