Collective attention in online social networks

Submitted by Tristan James Bowen Cann to the University of Exeter as a thesis for the Degree of Doctor of Philosophy in Computer Science, May 2021.

I certify that all material in this thesis which is not my own work has been identified and that any material that has previously been submitted and approved for the award of a degree by this or any other University has been acknowledged.

(Signature) .....

Abstract

Social media is an ever-present tool in modern society, and its widespread usage positions it as a valuable source of insights into society at large. The study of collective attention in particular is one application that benefits from the scale of social media data. In this thesis we will investigate how collective attention manifests on social media and how it can be understood. We approach this challenge from several perspectives across network and data science. We first focus on a period of increased media attention to climate change to see how robust the previously observed polarised structures are under a collective attention event. Our experiments will show that while the level of engagement with the climate change debate increases, there is little disruption to the existing polarised structure in the communication network. Understanding the climate media debate requires addressing a methodological concern about the most effective method for weighting bipartite network projections with respect to the accuracy of community detection. We test seven weighting schemes on constructed networks with known community structure and then use the preferred methodology we identify to study collective attention in the climate change debate on Twitter. Following on from this, we will investigate how collective attention changes over the course of a single event over a longer period, namely the COVID-19 pandemic. We measure how the disruption to in-person social interactions as a consequence of attempts to limit the spread of COVID-19 in England and Wales have affected social interaction patterns as they appear on Twitter. Using a dataset of tweets with location tags, we will see how the spatial attention to locations and collective attention to discussion topics are affected by social distancing and population movement restrictions in different stages of the pandemic. Finally we present a new analysis framework for collective attention events that allows di-

rect comparisons across different time and volume scales, such as those seen in the climate change and COVID-19 experiments. We demonstrate that this approach performs better than traditional approaches that rely on binning the timeseries at certain resolutions and comment on the mechanistic properties highlighted by our new methodology.

Acknowledgements

I'd first like to thank my supervisors Hywel Williams and Iain Weaver. Their patient support and sage advice developed the scholar I am today, and will remain with me for the rest of my career. With their guidance I've developed a respect for simplicity in approaches, a deep appreciation for the value of a good figure and more importantly an excellent primer in how to make them. I hope they feel as much pride as I do when I think of all we have achieved over the last few years.

The other members of the SEDA lab also deserve thanks for lending a friendly ear when I faced difficulties and patiently listening to my ideas whenever I needed a sounding board. Being a part of such an open and welcoming group during my PhD has been a pleasure.

My PhD was only possible thanks to an EPSRC studentship (grant number EP/M506527/1). I greatly appreciate all of the opportunities that this funding gave me.

My family deserve special recognition for the support they've given me over the course of this PhD. My parents have been an invaluable source of support and encouragement, especially following the disruption caused by the COVID-19 pandemic. This thesis is as much their achievement as it is mine.

My final thanks go to all my friends and housemates who constantly reminded me that there is more to life than my PhD. This would have been much harder without you all.

Contents

List of Figures

List of Tables

Author's declaration

Several of the chapters in this thesis have been previously published and contain work by my coauthors.

Chapter 4: Is it correct to project and detect? How weighting unipartite projections influences community detection

Chapter 4 has been published previously in two forms. The work, completed alongside Iain S. Weaver and Hywel T.P. Williams, was first published in the conference Complex Networks and the Applications VII in 2019 [33]. An extended version was published in the journal Network Science in 2020 [34], and it is this version that has been included in this thesis. For this paper I:

- Designed the experimental setup to compare different weighting schemes.
- Generated the experimental results across the different weighting schemes.
- Interpreted the results.
- Designed the network visualisations of the Twitter data.
- Wrote the first draft of the manuscript, and edited it in response to comments from my coauthors and reviewers.

The version printed in this thesis is as published in Network Science with minor edits for clarity following a change in reference styles.

Chapter 5: Ideological biases in social sharing of online information about climate change

Chapter 5 has been published in the journal PLOS ONE in 2021 [35] and was completed alongside Iain S. Weaver and Hywel T.P. Williams. For this paper I:

- Designed the experiment for temporal analysis of sharing structures.
- Analysed the Twitter data.
- Managed the team of manual coders when determining the domain level bias.
- Processed the article data.
- Interpreted the results from all experiments.
- Designed the network and bias visualisations.
- Wrote the first draft of the manuscript, and edited it in response to comments from my coauthors and reviewers.

The version printed in this thesis is as published in PLOS ONE and reproduced under the Creative Commons Attribution licence.

Chapter 6: A tale of two lockdowns: Comparing mobility and attention responses in two periods of coronavirus restrictions in England and Wales

This work has been conducted independently with discussion input from Rudy Arthur, Iain S. Weaver and Hywel T.P. Williams.

Chapter 7: Shapes of collective attention online

Chapter 7 contains work alongside Iain S. Weaver and Hywel T.P. Williams. This work has previously been submitted unsuccessfully for publication and subsequently improved following reviewer comments. For this paper I:

- Developed the methodology in the scale-independent representation.
- Performed the data processing and analysis steps.
- Interpreted the results.
- Wrote the first draft of the manuscript, and edited it in response to comments from my coauthors and reviewers.

Chapter 1

Introduction

Online social media platforms have emerged over the past 20 years and presented opportunities to connect with others across the world. By early 2019, 72% of US adults were using one or more social media platform such as Facebook or Twitter [3]. The resulting datasets covering millions of users can enhance many applications by improving coverage of groups or areas under-represented by traditional survey methods or sensor networks. The negative consequences of social media have also begun to emerge through accusations of electoral interference and the rise of misinformation in recent years. These different applications highlight the number of roles that social media can play in society. Social media can be a force for good, but it can also have negative impacts under the influence of malicious actors. Understanding and managing these divergent outcomes from social media will likely be critical for their continued contribution to modern life.

One factor that many aspects of social media share is a reliance on attention, in terms of both individuals and groups. The study of collective attention considers the latter process and aims to understand both the mobilisation of the attention of groups towards certain topics and the resulting patterns in social networks. World events can also catalyse collective attention and recognising why some events trigger far-reaching responses while others barely register among users is an important goal for the future development of social media platforms. Beyond the financial benefits to social media companies and advertisers, further knowledge of the dynamics of col-

lective attention will be invaluable for combating the rise in misinformation around a number of social and political issues through the early recognition of problematic content and effective disruption of spreading mechanisms.

This thesis will pose and answer a series of questions about how societal-level attention to events is reflected in social media behaviours at the individual and collective level. In the following chapter, we will give the reader an introduction to the concepts and techniques of network science and how they relate to social media. In addition, we will present an overview of how we define collective attention and how it has been studied in the past. Chapter 3 then outlines the questions we seek to answer and their importance. In Chapter 4 we address the first of these questions by testing a series of bipartite projection weightings to see which is preferred for community detection purposes. We then use this preferred methodology in Chapter 5 to study the change in attention as seen through community structure around a disruptive media event in the climate change discourse. Chapter 6 continues this exploration of disruptive events by looking at how the lockdown restrictions in response to the COVID-19 pandemic has affected the online social networks of England and Wales. In Chapter 7 we present a means of comparing events across different scales and show that this standardisation method also allows for characterisation of event periods. We finish by discussing the implications of these findings and possible directions of future research in Chapter 8.

Chapter 2

Background and literature review

Collective attention drives many interconnected systems in modern society. Advertising is a commonplace example of the power of collective attention, able to change awareness of brands or social issues and ultimately generate profits or political change. In this thesis we are interested in understanding how events cause collective attention to change. Social media is an important tool for this goal given its position as a regular means of communication for much of the world. The size of the user base is particularly advantageous as, along with the public nature of certain social media services, it means that large quantities of data are available to those interested in social systems.

There are some ambiguities in what is meant by collective attention. In many cases collective attention refers to the aggregate action of many individuals without considering how interactions affect this trend. An alternative view relies on this interaction. To give an example, consider a loud plane passing overhead. Many individuals may look up to investigate the noise and form a group paying attention to the same event. Now consider instead a similar situation, but this time without an auditory draw for attention. If one person is first drawn to the trail left by the passing plane, before others follow their gaze to the same object, we find ourselves in the second type of collective attention where the first individual's attention has influenced others. More concretely, this thesis considers collective attention to be the state where the attention of others increases an individual's attention. This

effect is much more difficult to measure, but where possible we prefer this definition going forward.

There may be differences in how collective attention behaves in offline and online spaces. The main contributing factor for this is likely to be the size of the group interacted with. On certain social media sites such as Twitter and Reddit, an individual can be in contact with millions of other users thanks to the public nature of most posts. Given the comparatively limited number of people that can be interacted with in offline contexts, collective attention events online are likely to involve many more people. Moreover, social media sites generate revenue from active user bases and therefore promote activity which can generate this interaction. These "trending" topics make it easier for users to become aware of the focus of others' attention and potentially promote a feedback loop feeding into further collective attention.

In this chapter we present the necessary background material for our investigation of collective attention, before highlighting how the literature has previously addressed questions in collective attention and related topics. We begin in Section 2.1 by giving a primer in the techniques of network science, and focus on those used to analyse structural patterns in large networks. Following this, Section 2.2 considers previous work on social networks, social media and collective attention through several lenses. The first lens looks at the field of computational social science more generally in Section 2.2.1. After this introduction we focus in turn on polarisation (Section 2.2.2) and information contagion (Section 2.2.3) as topics linked to collective attention, but often not explicitly stated as such. We conclude in Section 2.2.4 with a discussion of previous work to study collective attention through various means.

2.1 Network science

The field of network science has seen large growth in recent years after a series of methodological advancements and the availability of large datasets from online sources. Here we will discuss some preliminaries required for the subsequent investi-

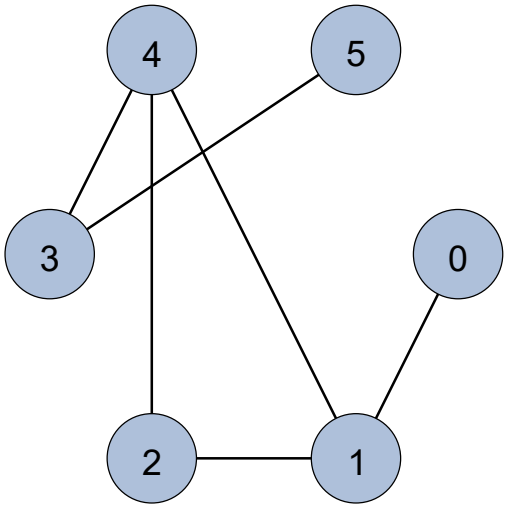
gations. An excellent source for further detail is Network Science by Barabási [19], which is freely available online[1].

Networks are representations of objects (usually described as *nodes* or *vertices*) and their relationships (usually described as *edges*). The literature uses many different terms for components of systems represented by networks and for simplicity we will use the following terminology for the rest of this chapter. In this thesis, we are primarily interested in social networks and therefore we call the participants in a system *actors* and they define the nodes of the associated network. Actors form *relationships* between one another which define the set of edges in the network. These representations are an efficient means of encoding relationships across many different experimental settings such as email interactions [100], scientific coauthorship [135], neural pathways of organisms [185] and drug-protein interactions in organisms [211]. Nodes and edges often have additional metadata associated with them to describe other relevant facets of the data. Edge weights are the most commonly used form of additional metadata and assign a real number to each edge to represent the strength of interaction. Where they exist edge weights are nearly always positive, although negative weights used to measure antagonistic relationships are not unheard of [102]. If edge weights are given, the network is said to be *weighted* and otherwise *unweighted*.

The mathematical study of networks builds on the existing frameworks developed in graph theory, and often uses an *adjacency matrix* to record the existence of edges between nodes, named as we use the term *adjacent* to describe a pair of nodes if an edge exists connecting them. Fig. 2.1 gives a small example network and its adjacency matrix. In this representation each row (and column) corresponds to a node in the network. In an unweighted network the adjacency matrix A is binary, $A_{ij} = 1$ signifies that an edge exists between nodes i and j, with no edge otherwise. In weighted networks the value A_{ij} gives the weight of the edge between nodes i and j. In the example in Fig. 2.1 we see that its adjacency matrix is symmetric, as we are not yet considering direction in the edges.

A number of attributes are used to describe networks in addition to the adjacency

[1] `http://networksciencebook.com/`



$$A = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \end{array}$$

Figure 2.1: A small sample network and its adjacency matrix. Row and column labels correspond to the nodes in the diagram.

matrix. The *degree* of a node is the number of edges that are attached to it. In the case of weighted networks, a node's *weighted degree* is the sum of the edge weights of all edges attached to the node. By collecting the degree of all nodes in a network it is possible to calculate the *degree distribution*, that is the probability of a randomly selected node having a specified degree. A *path* between two nodes is said to exist if a sequence of edges connect the two nodes through any number of intermediary nodes. Again, this can be extended to the case of the whole network and we say that a network in which a path exists between any pair of nodes is *connected*. We define the *components* of a network as the set of maximally connected node-disjoint subnetworks. The component with the most nodes is called the *giant component*.

Networks can be extended in many ways to represent different dynamics of the modelled system. A common extension is the notion of direction in the edges. In this way, the network is able to capture relationships which are not necessarily reciprocal. Some common examples of this are hyperlinks on webpages and following on Twitter. We call networks with a notion of direction in the edges *directed* and they enable the consideration of the edge $A \to B$ differently to the edge $B \to A$. In directed networks we define the *in-degree* (the number of edges pointing towards a node) and *out-degree* (the number of edges pointing away from a node) separately to better represent this directional behaviour.

2.1.1 Network models

Beyond their use as representations of real-world systems, network models exist to generate synthetic networks satisfying a series of desirable properties. Some commonly used network models emerged from the work of Paul Erdös, Alfréd Rényi, Albert-László Barabási and Réka Albert, who lend their names to the algorithms they devised.

The Erdös-Rényi model

At their simplest level, random network models create N nodes and then construct edges in one of two different ways. If a fixed number of edges L is required then these can be assigned randomly between all possible pairs of nodes, often referred to as a $G(N, L)$ network. Alternatively, edges can be defined in a probabilistic manner with an edge existing between each pair of nodes with some given probability p, analogously referred to as a $G(N, p)$ network. As the size of the network grows, these models become equivalent given appropriate choices for L and p, although the probabilistic edge definition gives no guarantee of a certain number of edges in the final network [19].

Given the simplicity of these two models, many of their characteristics can be derived analytically. In the case of $G(N, p)$ networks we can calculate the expected number of edges as follows. This network has N nodes, and therefore $N(N-1)/2$ possible node pairs exists (assuming that self-loops are impossible). An edge exists between each such pair independently with fixed probability p, and therefore the expected number of edges in such networks is $pN(N-1)/2$. Using a similar argument, we can see that the average degree of a node in a $G(N, p)$ network is $p(N-1)$. At this stage, the properties of the Erdös-Rényi model appear to be similar to that of a binomial distribution. In fact, we can see that the degree distribution of such networks is defined by a binomial.

Suppose we wish to calculate the probability that a node has degree k. Then exactly k of its $N-1$ possible edges exist, which happens with probability $p^k(1-p)^{N-k-1}$. Additionally, we do not require any specific edges to exist, so any choice of k from

the $N-1$ is valid. There are $\binom{N-1}{k}$ such selections. Therefore the probability of a node x having degree k is

$$P(deg(x) = k) = \binom{N-1}{k} p^k (1-p)^{N-k-1}, \tag{2.1}$$

that is the classical definition of a binomial distribution [19].

Preferential attachment and scale-free networks

The Erdös-Rényi model has a number of limitations for representing networks in the real world. These limitations emerge in the degree distribution and other statistical features of the network as a result of the formation process considered [19]. Instead of forming as a single entity with all nodes arriving at the same time, real networks grow with nodes and edges joining the system at different points in time.

The first addition required for networks to represent this real-world behaviour is a concept of growth, that is nodes join the system over time. With this additional complication, the creation of such a random network continues for a desired number of timesteps with each adding a new node and a number of edges. The edges from the new node can use either of the two approaches described for the Erdös-Rényi networks, but the growth aspect leads to subtle changes in the network characteristics. Foremost is a skewing in the degrees of nodes in the network, as nodes introduced earlier in the network's evolution have more chance to share edges with other nodes. This model has been studied in detail by Callaway et al. [32], who show that they are expected to have an exponential degree distribution.

Until this point, we have considered nodes as statistically equivalent when generating our random networks. This equivalence assumes that any node is equally likely to be selected when a new edge is to be added under growth, but this is counter-intuitive in many real-world situations. Consider the network of scientific citations in a particular field. While any of the papers in the field may be suitable for citation, those that already have a number of citations enjoy increased visibility and endorsement from other scholars. This concept of choosing high degree nodes when adding new edges became known as *preferential attachment* [18], and became

a core component of the Barabási-Albert random network model.

The Barabási-Albert model typically uses a linear weighting to preference high degree nodes when adding new edges. In most formulations, a new node is added at each timestep, with a fixed number of edges m being added between the new node and existing nodes in the network. The probability of a node i of degree k_i being selected is then given by

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j}. \tag{2.2}$$

If we take an initial network with n_0 nodes and e_0 edges and then progress for t timesteps we will have a network with $n_0 + t$ nodes and $e_0 + mt$ edges. It remains to be seen how the preferential attachment phenomena has affected the network characteristics. In his book [19], Barabási begins by calculating the dynamics of a single node's degree. The expected change of node i's degree when the Nth node is added can be expressed as

$$\frac{dk_i}{dt} = m\Pi(k_i) = m\frac{k_i}{\sum_j^{N-1} k_j} \tag{2.3}$$

since each of the m new edges has probability $\Pi(k_i)$ of connecting to node i. This statement can be simplified further. As we have already seen, the network at time t has $e_0 + mt$ edges, and each edge increases the total degree by 2. From this it follows that $\sum_j^{N-1} k_j = 2(e_0 + mt) - m$ and therefore

$$\frac{dk_i}{dt} = m\frac{k_i}{2(e_0 + mt) - m} = m\frac{k_i}{2e_0 + m(2t - 1)}. \tag{2.4}$$

From here we consider the regime of large t, which allows the $2e_0$ and -1 terms in the denominator to be ignored, further simplifying the rate of degree change to

$$\frac{dk_i}{k_i} = \frac{1}{2}\frac{dt}{t_i}. \tag{2.5}$$

We can solve this equation using the fact that $k_i(i) = m$, that is a newly added node has degree m to find

$$k_i(t) = m\left(\frac{t}{t_i}\right)^{\frac{1}{2}}. \tag{2.6}$$

In more generalised cases, the exponent is labelled β and called the *dynamical exponent*. Expressing the degree of a node at time t allows us to now calculate the degree distribution under the assumption of stability as we enter the scale-free regime for sufficiently large t [20]. Our first goal is to express the number of nodes with degree smaller than k at time t. Using Equation 2.6 we have

$$m\left(\frac{t}{t_i}\right)^{\frac{1}{2}} < k \tag{2.7}$$

which we rearrange to

$$t_i > t\left(\frac{m}{k}\right)^{1/\beta} \tag{2.8}$$

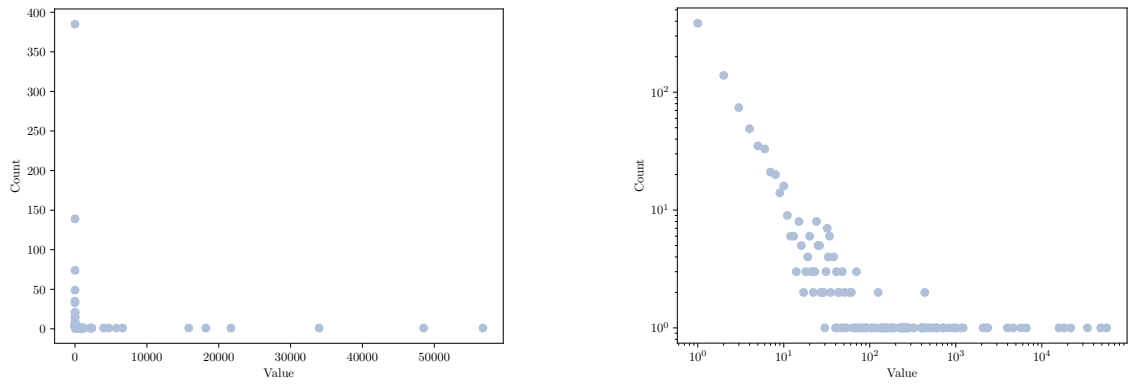
where $\beta = \frac{1}{2}$. In this representation, t_i is the timestep at which node i was added and since nodes are added at a constant rate, the number of nodes satisfying this equation is given by the right-hand side of Equation 2.8. Using this number we can estimate the proportion of all nodes satisfying the criteria and hence estimate the probability. At time t, we have already seen that there are $n_0 + t$ nodes in the network. For simplicity we consider the regime of large t where the contribution of n_0 becomes insignificant and use t to approximate the number of nodes in the network. Therefore the cumulative degree distribution $P(k)$ becomes

$$P(k) = \frac{1}{t}\left(t - t\left(\frac{m}{k}\right)^{1/\beta}\right) = 1 - \left(\frac{m}{k}\right)^{1/\beta}. \tag{2.9}$$

All that remains to recover the degree distribution it to differentiate Equation 2.9 with respect to k, which gives

$$p_k = \frac{1}{\beta}\frac{m^{1/\beta}}{k^{1/\beta+1}} = 2m^2k^{-3}. \tag{2.10}$$

Equation 2.10 shows that we have reached a special regime of networks, where each degree is determined by a so-called *power law distribution* where the negative



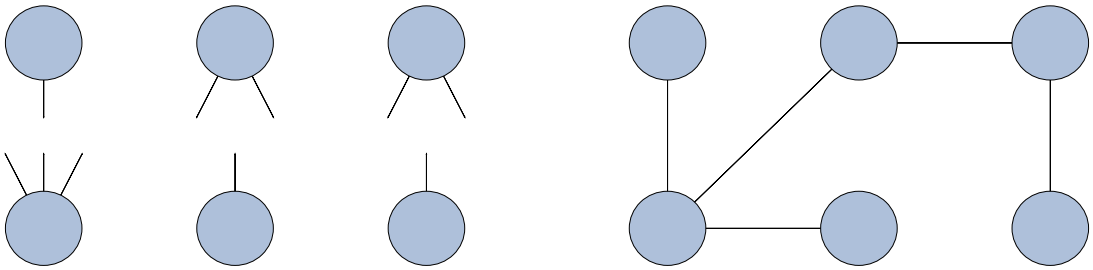
(a) Sampled values on linear axes.

(b) Sampled values on logarithmic axes.

Figure 2.2: An example of a zeta distribution with exponent $\frac{3}{2}$. Note the characteristic straight line for much of the sample when plotted on log-log axes.

exponent of k (often labelled γ) is called the *degree exponent*. In practical contexts, node degrees are necessarily integers, and as such can be equally represented by the zeta distribution, a discretised complement to the power law distribution. Such networks are also called *scale-free*. Networks with scale-free characteristics are common in the real-world, including examples such as website hyperlinks and citation networks (e.g. [6, 87]). Another important example is the directed network of followership on platforms like Twitter with no guarantee of reciprocity. In this case, both the in-degree (number of followers) and out-degree (number of users followed) both displayed heavy-tailed degree distributions, but were best fit by power law and log-normal distributions respectively [133]. This example highlights an important characteristic of directed networks - the in-degree and out-degree can display differences in degree distribution.

Fig. 2.2 shows an example of values sampled from a zeta distribution with exponent $\frac{3}{2}$. Visualising on linear axes (Fig. 2.2a) clearly shows the typical patterns of scale-free distributions with many small values, and a few very large values. Plotting the same distribution on log-log axes (Fig. 2.2b) illustrates how it can be used to give an intuition of whether a given degree distribution likely belongs to a scale-free network.



(a) A series of edge stubs are generated.

(b) The stubs are then matched at random.

Figure 2.3: The configuration model generates random networks by sampling node degrees from the required degree distribution. (a) Each node has a number of edge stubs created to match its degree. (b) Edge stubs are paired at random to create the final network.

Random networks in practice

In the previous section we have seen the mathematical derivation of several common random network models and developed a better understanding of the network development processes they represent. The downside of the models as presented here is they can be time consuming to run for large network sizes, but fortunately there are processes to alleviate this issue.

The first approach is know as the configuration model [206]. The configuration model randomly samples a network from a given degree sequence using a series of constructive steps. From this degree sequence, a series of edge stubs are created across all nodes in the network. Finally, the edge stubs are joined at random. For a visual outline of this process, see Fig. 2.3. In practical contexts the configuration model can be used with a desired degree distribution given a series of sampled degree sequences.

There are a couple of minor complications to consider when working with the configuration model. Firstly, the sequence of node degrees must have an even sum to avoid any unconnected stubs. In practical situations with large networks, this is not usually a problem as the redundant stub can be deleted without much impact on the characteristics of the final network. The other main issue is around the number of components. With random stub matching there is no guarantee that the generated network will consist of a single component. How significant an issue this is

depends on the system and circumstances represented, but again in large networks it is possible to ignore. In such cases, the giant component can be of comparable size to the whole network, given suitable simulation conditions [141].

The second means of generating random networks in practice steps away from the abstract models we have seen before. Using an existing network, it is possible to generate new networks with the same degree distribution using a process of rewiring. In a similar process to the configuration model, the edges in the existing network are broken into a series of edge stubs. As before, these edges stubs are then paired up at random to construct a new network.

Random rewiring avoids the parity problem of the configuration model, but is still susceptible to fragmentation of the network into multiple parts. One major advantage with this technique is that it presents a quick and effective means of simulating a null model for different characteristics. Using an ensemble of randomisations allows researchers to measure significance of network properties compared to a particular degree distribution. This technique has seen varied uses from estimating diversity in Airbnb users [101] to measuring the existence of structural patterns [68].

2.1.2 Common network metrics

As large and complex objects, networks often require summarising to be interpreted and compared. Many suitable metrics have been proposed over the years, and we choose to focus on some of the most common here.

Betweenness and centrality measures

The first measure we examine is called *betweenness*. Betweenness arises from the need to identify which nodes in a network are most important for following paths of edges through a network. The typical measure for this at the node level is given by

$$\mathrm{Bet}(n) = \sum_{s,t \neq n} \frac{|\mathrm{Paths}_n(s,t)|}{|\mathrm{Paths}(s,t)|}, \tag{2.11}$$

where $\mathrm{Paths}(s,t)$ is the set of all paths between nodes s and t, and $\mathrm{Paths}_n(s,t)$ is the set of such paths that pass through n. In other words, betweenness measures the

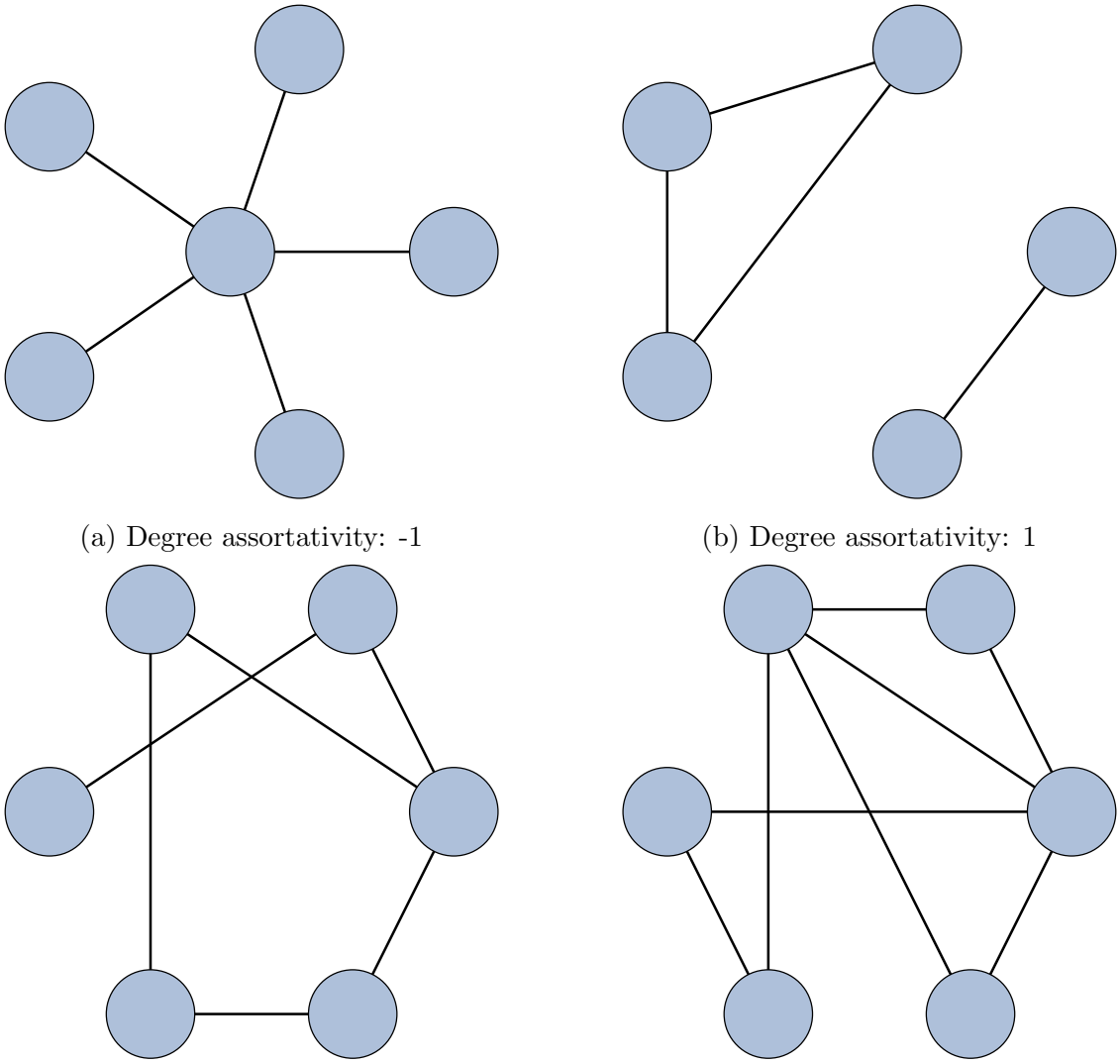
proportion of all paths in the network that pass through the given node. In practical contexts, this gives a measure of each node's importance for the relationship modelled by the network, for example nodes with a high betweenness in a social network are those who facilitate much of the wider social contact among the participants.

Betweenness is one of many so-called *centrality* measures used to identify important nodes in a network based on certain characteristics. We've seen that betweenness centrality identifies as important those nodes which facilitate many paths in the network. Other measures instead focus on number of incident edges (degree centrality, simply each nodes degree), nodes closer to all other nodes in the network (closeness centrality, average shortest path length from a node to all others in the network) or influence of nodes within the network (eigenvector centrality, typically preferring nodes connected to other central nodes [8]). Other more complex centrality methods have also seen use in different contexts, most notably the PageRank centrality [146] used to power the Google search engine at the beginning of its lifespan.

Assortativity

We now transition from looking at properties of individual nodes as we did with centrality to properties of the network of a whole. A key property at this scale is *assortativity*, a measure of preference for edges in the network to connect nodes that are similar in some way. A common way to measure this is to calculate the Pearson correlation coefficient of node degrees for each pair of connected nodes. In this context, a high correlation means that high degree nodes are more likely to connect to other high degree nodes, and vice versa [136].

Assortativity is illustrative of different aspects of network structure. In networks with low assortativity, nodes with low degree generally connect to nodes with high degree. This network structure is often referred to as hub and spoke, and can represent situations such as technological networks where many users connect to the same central services. At the other end of the scale, nodes in networks with high assortativity tend to form edges with nodes of similar degrees such as social



(a) Degree assortativity: -1

(b) Degree assortativity: 1

(c) Degree assortativity: -0.09

(d) Degree assortativity: -0.5

Figure 2.4: Examples networks and their assortativity scores.

networks [136]. Some example networks and their assortativity scores are given in Fig. 2.4.

Clustering

Moving away from micro-scale measures like centrality that consider individual nodes and the macro-scale measures like assortativity that consider the whole network, we can instead look at *clustering*, a metric that can be applied between these two extremes. Clustering is a measure of local density in the network, and looks to measure the existence of triads around a node. More specifically, for a node n its clustering coefficient is

$$C_n = \frac{2|\{e_{ij} : \exists\ e_{ni} \text{ and } e_{nj}\ \forall\ i, j \in N(n)\}|}{k_n(k_n - 1)}, \tag{2.12}$$

where $N(n)$ is the neighbourhood of n, that is the set of all nodes adjacent to n [19]. More concretely, we measure the proportion all possible edges that exist between all pairs of neighbours of node n. The definition in Equation 2.12 applies to undirected networks, but without the factor of 2 applies to directed networks. Following this calculation of clustering coefficients at the node level, it is possible to derive a measure at the level of groups of nodes, or the entire network, simply by finding the average node clustering coefficient over the set of nodes [19].

2.1.3 Modularity and community detection

The clustering coefficient outlined in the previous section gives a measure of how densely interconnected the network is, but gives no indication of how these structures fit within the entire network. This goal requires community detection, that is the process of identifying groups of nodes that are more likely to connect to one another than another node selected at random. Defining an algorithm to calculate such a community structure faces difficulties in complexity and a quality metric. In terms of complexity, community detection has similarities to graph partitioning, that is the process of dividing the nodes in a graph into disjoint subsets. This problem is considered to be NP-complete, and can only be solved in polynomial time using heuristics and special cases (e.g. the Kernighan-Lin algorithm for sparse graphs [98]). The choice of potential metrics introduces a similar difficulty, as many different aims must be balanced. Often this will involve the definition of a null model to compare the structure against and ask whether the observed structure is different compared to that which would be expected in a random network. As we saw in Section 2.1.1, many different characteristics can be incorporated to subtly change the model's behaviour.

Over the years, several approaches have been proposed (e.g. betweenness [140], clique conductance [120]) but one family of methods has grown in popularity to become the most commonly used. Modularity-based methods, named after the

community quality metric developed by Newman and Girvan [140], provide an efficient measure of community quality at each step of an algorithm by comparing to a null model that simulates a degree-preserving rewiring of all edges in the network. Using this null model, modularity compares each pair of nodes in the network with the null model and preferentially scoring stronger within community connections than expected and penalising weaker ties than expected. More concretely for each node pair i, j, modularity calculates

$$\left(A_{ij} - \frac{k_i k_j}{2E}\right) \delta(C(i), C(j)), \tag{2.13}$$

where A_{ij} is the i, jth element of the network adjacency matrix A, E is the number of edges in the network, $C(i)$ gives the community label for node i and $\delta(a, b)$ returns 1 when $a = b$ and 0 otherwise. Averaging this over all node pairs in the network we get

$$Q = \frac{1}{2E} \sum_{i,j} \left(A_{ij} - \frac{k_i k_j}{2E}\right) \delta(C(i), C(j)), \tag{2.14}$$

the modularity of the community assignment on the network, first detailed in [140]. Modularity values fall within the range $[-1, 1]$, where negative values indicate that there are fewer edges within communities than expected at random, positive values indicate the communities contain more edges than expected at random and 0 shows that the edges within communities are as expected under the rewiring model. In practice, a given community structure is considered to be significant if the modularity is greater than 0.3 [42].

Community detection algorithms use modularity as a metric to perform a series of small changes to the current community assignment and improve the current modularity score. The simplest such approach is a greedy algorithm first proposed by Newman [137] which begins by starting each node in its own community before iteratively merging the communities that give the largest increase in modularity. While this greedy algorithm presents no guarantee of finding the optimal community configuration, its sacrifice in accuracy presents savings in efficiency for large

networks.

In the search for improved modularity maximising algorithms, Blondel et al. [24] propose the Louvain method which builds upon the basic framework of the greedy method to provide additional context to the communities. The Louvain algorithm divides the community detection into a series of iterations, each comprised of two phases. In the first phase, each node is assigned to its own community, and then successively tested to see if removing it from its current community and adding it to the community of one of its neighbours will increase modularity. This phase continues until no increases in modularity are possible, at which point each of the communities are induced into a single node in a new network. These iterations are repeated on the new network until no more modularity increases are possible. At this stage a tree of community structure can be output to visualise the tiered structure of the communities. The Louvain algorithm also benefits from increased efficiency thanks to the ease with which the change in modularity for a single node can be computed [24].

Community detection using modularity does suffer from a notable drawback however. Modularity maximisation experiences what is known as a resolution limit, where small communities are merged under certain circumstances. If the total degree in a pair of communities is sufficiently small compared to the number of edges in the network, then merging these communities will increase modularity if even a single edge exists between nodes in the two communities [69]. As a result, some communities may in fact be a coalescence of many smaller communities. This potential drawback has not seen modularity maximisation lose favour in the scientific community, but does highlight the need to consider communities carefully in the context of the network and the situation it represents.

Beyond modularity maximisation, stochastic block models are another method of community detection that has grown in popularity in recent years. Stochastic block model methods detect community structure in networks by fitting a random network model characterised by a series of communities where the probabilities of edges existing between nodes are based only on their community assignment. As a result

they provide a clear indication of the statistical significance of any detected structures. Such methods avoid the resolution limit seen with modularity maximisation since they allow for variable community sizes and preferences between them (and in fact have been shown to be equivalent to modularity maximisation if the communities are treated as statistically equivalent [139]). Adaptations to this approach have been made to allow for detecting hierarchical structures [148], distinguish between assortative and disassortative community structures [119] and recognise whether community structure is formed by truly homophilous relationships or is an artefact of triadic closure [149]. While these features are advantages of stochastic block models over modularity maximisation, these tools are not yet included in widely used software packages. As such, this thesis will focus on modularity maximisation for community detection.

2.1.4 Bipartite networks

The definitions and metrics explained so far in this section refer to the simplest type of network, where edges are allowed between any pair of nodes. However there are many cases where such behaviour does not truly represent the system to be studied. In these cases the nodes can be divided into two or more classes (or modes) and it is impossible for edges to exist between nodes in the same class. The vast majority of examples use just two classes and are called *bipartite* networks, and the two classes are often referred to as left and right, or top and bottom. Examples of bipartite networks that have been studied in the past include actors and films they appear in [107], scientists and the papers they publish [25] and people and the events they attend [58].

On the face of it, this extra condition does not change many of the aspects of networks that we have seen before. Bipartite networks can still be weighted or unweighted, directed or undirected and each node still has a well-defined degree. The adjacency matrix can be expressed in a more succinct way, where the rows correspond to the left nodes and the columns correspond to the right nodes (since a representation with a row and column for each node has blocks of zeros along

the diagonal). The centrality and assortativity metrics we saw in Section 2.1.2 also extend without change to the case of bipartite networks.

But while some characteristics extend without major changes, there are others that require updating to reflect the new edge restriction imposed by the two bipartite modes. The degree distribution of bipartite networks is not necessarily consistent across both modes, and therefore it is often important to examine the degree distributions of the left and right modes separately. Clustering also requires adjustment. The measure we see in Section 2.1.2 is looking for triangles in the network, i.e. sets of three mutually connected nodes. Triangles cannot appear in bipartite networks, since they would require a within-mode edge. Consequently, the notion of clustering must move away from such a focus. Latapy et al. [107] propose an intuitive extension that consider nodes to be highly clustered if for two nodes in the same mode there is a large overlap between their neighbourhoods. More concretely, they define the bipartite clustering coefficient to be

$$C_{u,v} = \frac{|N(u) \cap N(v)|}{|N(u) \cup N(v)|}. \tag{2.15}$$

To define this for a single node n, the paired clustering coefficient can be averaged over all nodes with which n shares a neighbour, that is

$$C_n = \frac{\sum_{v \in N(N(n))} C_{n,v}}{|N(N(n))|}, \tag{2.16}$$

and a similar averaging procedure can be used to calculate the clustering coefficient over one of the modes, or the entire graph.

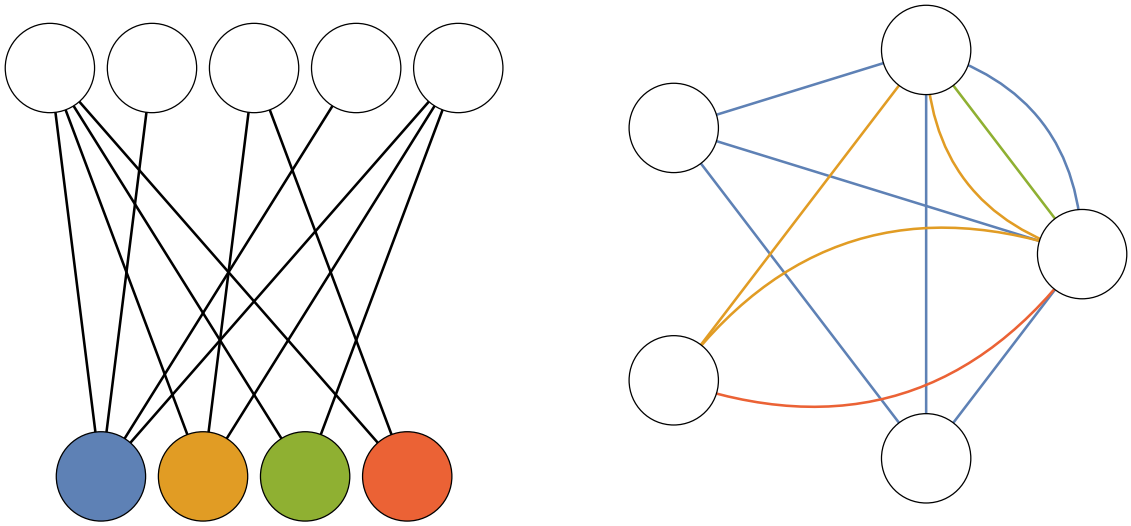
Random network models also require adaptation to reflect the edge restriction of bipartite networks. The easiest way to do this in practice is to extend the concept of the configuration model, generating a degree distribution for each mode in the network, before connecting the edge stubs in an appropriate way. As before, this presents the potential issue of imbalanced degree sequences, but, as before, this problem is insignificant in sufficiently large networks. This approach does however allow for easily varying the degree distributions for each mode, which can be desirable in certain circumstances.

Another major difference for bipartite networks is seen in modularity-based community detection. The requirement that edges in bipartite networks only exist between nodes in different modes is not adequately reflected in the null model incorporated in the modularity calculation described in Section 2.1.3. Attempts have been made to incorporate alternative null models that reflect this restriction, the most commonly used is that of Barber [21] which adapts a null model of a degree preserving rewiring of the network while reflecting the two-mode split in the network.

In contrast to modularity maximisation approaches, stochastic block models methods for community detection are more readily applicable to bipartite networks. Choosing community blocks and edge probabilities that reflect the edge restrictions of bipartite networks allows stochastic block model approaches to detect communities in bipartite networks with no further modifications. This has been shown to be an effective technique [105], but as mentioned previously stochastic block model tools are not available in many widely used software packages.

One method to circumvent these difficulties with bipartite networks is *unipartite projection*, a process which infers a unipartite network with edges between the nodes in one of the modes in the bipartite network. Edges in the unipartite projection exist between nodes that share neighbours in the bipartite network, and thus represent the "hidden" relationships between nodes in one mode of a bipartite network. In this way it can be seen that many unipartite networks studied are implicitly projections of some underlying bipartite network. A classic example of this is a coauthorship network, which is a projection of an author-article bipartite network.

Unipartite projections have drawbacks of their own that require special consideration in their analysis. Unipartite projections are formed of the superposition of cliques, as shown in Fig. 2.5. Nodes in the projected out bipartite mode contribute cliques of size equal to their degree in the projection. This clique composition has implications for many of the characteristics of the unipartite projection. While efforts to understand the impact of the projection process on various network characteristics have been studied in some circumstances (e.g. degree distributions [181] and network metrics [113]), some interactions remain understudied. A particular



(a) A bipartite network between coloured and uncoloured nodes.

(b) Each coloured node in the bipartite network produces a clique in the unipartite projection.

Figure 2.5: The unipartite projection process infers a new network from the composition of cliques for each projected out node (identified by edge colour). Edge weightings can be applied by summing repeated edges, or other more sophisticated methods.

challenge is the consequence of high degree nodes producing large cliques and increased clustering in the projection compared to a random network with the same degree distribution. One solution to this issue is appropriate weighting of edges in the projection, and we will see this in more detail in Chapter 4, which looks at how projection weighting can improve the accuracy of community detection. Alternatively, as with Barber's approach, the null model of modularity can be adapted to reflect this behaviour (e.g. [13]).

2.2 Networks, social systems and society

Following the brief introduction to network science in the previous section, we are now in a position to explore how scholars have made use of the available tools and techniques to get a deeper understanding of the evolution of social systems and how they interact with society. In this section, we will initially explore the broad concept of computational social science before focusing on the topics of polarisation, information spreading and attention propagation. We prioritise these three areas as important components of the processes driving collective attention (in the

case of information spreading and attention propagation) and a frequent structural component of networks under collective attention to certain topics (in the case of polarisation).

2.2.1 Computational social science

Argued for as a field in its own right in 2009 by Lazer et al. [108], computational social science has emerged alongside the growth of computing power and datasets, typically focused on how big data and computational methods offer new opportunities to tackle questions about how people interact and form social systems. The growth of online social networks since the early 2000s has had a particularly large impact on the field by making available datasets of millions of participants in largely natural situations.

Work by researchers in this area is typically split into two approaches. In many cases, the first approach is to use computational methods to perform experiments on data gathered from various sources. Alternatively, modelling studies develop theoretical explanations for the various behaviours observed in their situations of interest, often probabilistic or requiring large scale simulation, before comparing these modelled behaviours against those observed in the real world.

Social networks

A natural approach for studying the relationships between entities is the construction of social networks. Over the years, scholars have studied the relationships between individuals or organisations in applications as diverse as scientists and their citation networks [201] or companies and their board member networks [141]. In any case, the goal is often to use the patterns in how different participants in the system interact to reveal nuances that may be otherwise difficult to observe.

Two common features studied in social networks reflect how new edges are formed. Homophily, a generalisation of assortativity, expresses the preference for individuals to connect with others who are like them in some way. In some sense, homophily can be thought of as a bias in the network, and has been observed as

such in several studies. Koh et al. [101] measure homophily in rentals on the Airbnb platform, which in this case manifested as a preference for renters to choose properties posted by individuals like themselves along some demographic dimension. Homophily is in many cases linked with triadic closure, a measure which expresses how likely two individuals with a mutual contact are to be connected themselves. Melamed et al. [125] studied the influence of network structure on the cooperation of individuals under different levels of knowledge about other participants. Increased triadic closure, in this case indicating greater cooperation, was found to occur when a reputation for cooperation among prospective collaborators was made available. Given these examples, we can see how homophily and triadic closure are key measures of how important preference is in the development of a given social network.

Community structure is also an important means of understanding the properties of social networks. Algorithmic community detection such as through modularity maximisation can provide information about social systems by revealing participants that interact more frequently. In some cases, communities recover known properties of the actors in the social network. By considering coauthorship of scientific papers, Newman and Girvan [140] were able to recover university groups, subject areas and subfields shared by different scholars. Community structure does not always recover the expected behaviour however. Expert et al. [59] did not recover the linguistic split into French and Dutch speaking regions of Belgium from call records using modularity maximisation. In their case, the structure was recovered using a method that reflected the geographic distance between points, but in other contexts these differences should be investigated to see if unknown underlying factors are responsible.

Social networks and social media

The study of social media has grown in tandem with the study of social networks. The tools of network science are often deployed to analyse a necessarily connected system of many individuals. Moreover, the scale of social media data has driven the development of efficient new techniques to aid interpretation of large datasets.

Social networks constructed from social media data can equally take unipartite or bipartite forms. When represented as unipartite networks, social media data is typically modelled through the interactions of users. Depending on the social media site in question, these interactions can vary greatly. Followership or friendship of individuals has seen much investigation (e.g. [45, 86, 194, 196]), as have reposting behaviours (e.g. [74, 78]). Bipartite network representations extend the scope of the networks considered to include interactions between users and content. In the context of most social media services, this frequently covers hashtags, a means for users to tag topics as relevant to their posts. Such formulations can track the trending topics of conversation around events [26]. Alternatively, network representations have considered the relevance of pages to users either through their user activity [164], or the links they share [197]. Given the challenges in working with bipartite networks, projection is often used either implicitly or explicitly (e.g. [52, 71]). Regardless of the approach taken to analyse the network, each of these studies uses the additional structure of the bipartite formulation to investigate new explanations of the patterns seen in social media data.

Social media data is particularly advantageous for studying the structure of social networks as the volume of data available makes it possible to consider how communities change over time. Lorenz-Spreen et al. [118] included a dynamic component to their study of communities defining hashtag topics by tracking their evolution over time. Weaver et al. [188] adopted a similar approach to measure the shifting allegiances between politicians in the UK by linking communities between the temporal layers in a multiplex network. One major challenge with such approaches is the introduction or removal of nodes between steps and as such this technique is more easily applied to contexts with more stable node sets, such as elected politicians.

2.2.2 Polarisation and echo chambers

One area that has particularly benefited from the availability of big data is the study of political opinions among the public. Previous approaches were reliant on survey methodologies or controlled experimental settings, which may have issues

with participant honesty (in the case of surveys) or natural behaviours (in the case of controlled settings). Of particular note are studies exploring the differences between how left- and right-wing groups operate in society. Such work often considers the phenomenon of polarisation. Polarisation can be used to describe both a structural behaviour where there is limited contact between disparate political beliefs, or the process that causes such patterns.

Much work has sought to understand how differences in communication strategies have influenced the perceived growth in polarisation in recent years. Freelon et al. [70] give a good overview of the campaigning strategies of political activist groups and their expected impact, whereas Conover et al. [46] give an in depth analysis of how user activity on Twitter differs by political affiliation. Before the turn of the millennium, polarisation was believed to be stable overall but is increasing under certain issues [53]. The increase in polarisation has been measured more recently around the topic of climate change [54]. Through the use of social media data, this pattern has been studied in many contexts. Conover et al. [45] found divisions between left- and right-leaning individuals in their direct interactions, although the strength of polarisation was dependent on the type of interaction considered, a finding echoed by Williams et al. [196] in the context of climate change. In these two cases, community structure in various network constructions were used to measure polarisation. Other contexts have made use of homophily (e.g. [43]) as an alternative network feature. While many of these studies have been focused on English-language networks, likely due to their convenience for any subsequent text analysis, the broad trends have been similarly revealed in other linguistic settings e.g. Switzerland [72], Egypt [27] and Germany [82].

One main area of focus that has developed over the past five years is the impact of media interactions on polarisation. Fuelled by perceived changes around the Brexit referendum in the UK, the US presidential election in 2016 and associated revelations about targeted advertising, increased effort has been devoted to exploring how media followership and information engagement have affected interaction between different political beliefs. Numerous studies have observed the

reported polarisation online in the context of how users engage with politicised content (e.g. [6, 52, 93, 164, 172, 187, 197]). This pattern is typically revealed in content that users have chosen to share as the clearest available indicator of personal awareness and engagement. These concepts of awareness and engagement pose specific challenges for scholars in terms of accessing the true extent of information viewed in authentic settings. One possible solution to this issue is to gather all of the internet use data for individuals over a period of time. Gathering such data from private homes carries with it a number of privacy and ethical concerns, which have thus far been impossible to satisfy. Monitoring such behaviour in laboratory conditions with informed consent will capture many of the desired interactions, but there will always be the possibility of self-moderation by the participants affecting the trends observed. Therefore despite the potential for missed interactions, content included in public posts gives the clearest signal currently available.

Beyond the trends observed in interaction with content, scholars have investigated how the choice of which accounts to follow affects a user's place in the social network. Colleoni et al. [43] use homophily to compare the polarised structures formed around users on Twitter, alongside a classifier to determine a user's affiliation with the Democrat and Republican parties in the US. Using "active" and "passive" affiliations to distinguish between followers of party accounts and labels from the classifier, they found that passive Democrats exhibited less homophily, while active Republicans demonstrate greater homophily. A similar result was derived from followership of media accounts on Twitter around the 2017 "Unite the right" rally in the US. Tien et al. [178] used principal component analysis on followership of a small group of media accounts as an effective way to estimate user political views. This pattern was shown to extend beyond party political views and also affect opinions on issues such as climate change [65].

A closely related concept to polarisation, echo chambers are another often studied characteristic of how large groups interact. Echo chambers are a structural phenomenon in social networks, wherein participants are repeatedly exposed to the same information or opinions. In this way, it is easy to think of echo chambers as

forming from two components, a "chamber" of densely connected participants and an "echo" which persists in the group [94]. As with polarisation, echo chambers have been identified in direct interactions [196], news sharing [190] and inferred interactions on policy decisions [94]. In addition to observational studies, scholars such as Bright [29] have tried to link the formation of echo chambers to different levels of underlying ideology and extremism, whereas Garimella et al. [74] look at controversial topics and how they influence the formation of echo chambers.

The examples given above focus on topics that are becoming more politicised over time [53]. Through this politicisation, it becomes natural to question whether these structures are arising naturally or are artificially reinforced by the choices of the users. Here we draw in the concepts of selective exposure (also known as confirmation bias). Selective exposure suggests that users prefer to seek out information that agrees with their existing internal opinions and biases [15]. Much work has explored how user choices shape the development of polarised structures. Himelboim et al. [86] manually code communities of political tweets, and provide more evidence that the more politicised a topic becomes, the greater the level of polarisation it experiences. Other work has drawn on the difference between information that is actively sought out and that which is passively displayed. Weeks et al. [190] found that this incidental exposure led to more selective exposure by users. This finding was supported in the area of climate change polarisation by the work of Bakshy et al. [15] who showed that selection bias had a larger impact than algorithmic effects. Earlier work by Messing and Westwood [127] tested a similar hypothesis in a laboratory setting by recording article selection under different social cues. They found social endorsement led to greater counter-attitudinal source selection, which was not the case when exposed to the source alone. Considering all of this evidence together therefore presents a strong argument that selective exposure is driving the formation of the widely-recognised polarised structures online.

Selection bias cannot be separated from homophily or the existence of echo chambers in social networks. If selection bias truly exists among the participants in a system, then homophily should be expected as new edges are more likely to be

formed between similar individuals. Taking this a step further, at the scale of the whole network any selection bias is likely to develop echo chambers through aggregation of homophilous relationships. In light of this, analysis of any of these processes should be mindful of the relationship between them.

2.2.3 Spreading dynamics of (mis)information

The previous section looked at how opinions are reflected in social network structure, and the part that information plays in these opinions. This is not the only aspect of information use that has been studied in the context of social media. The influence of social dynamics on the spreading of information has grown alongside the phenomenon of virality, the notion that certain content can spread rapidly across large groups and distances through social interaction.

Given the name virality, this notion of information spread is commonly linked to the classical epidemiological models used for various pathogens. A frequent starting point for such work is the susceptible-infected-recovered (SIR) model [97]. The SIR model assigns each member of the population to one of three states: susceptible, i.e. those not yet exposed to the contagion; infected, i.e. those that are currently carrying the contagion and capable of passing it on; and recovered, i.e. those that are no longer receptive to the contagion and unable to pass it on. Here the parallels between diseases and information are a little clearer as new information and diseases are able to be known or contracted by new people. As the contagion progresses, sick people recover or information loses relevance and the population becomes unaffected by subsequent exposure to the contagion. The SIR model usually considers each stage in turn, with a proportion of each class advancing to the next stage with characteristics determined by the type of contagion modelled. More sophisticated modelling efforts incorporate contact networks to more realistically represent the role of interpersonal connections on transmission. In these cases, susceptible people have a chance to be infected for each infected neighbour they have, while infected people have a chance to spontaneously transition to the recovered class. Simulation of such setups can provide a better understanding of a particular contagion, and

inform public health decisions [150].

As is often the case with studies of disease, analysis of information spreading is typically concerned with which population characteristics affect the speed and scale of the diffusion. Kitsak et al. [100] applied the SIR model to email contact, inpatient interaction and film costar networks as specimen social situations. By varying the number and location of the initially infected nodes, they were able to identify links between node degree, degree-based structures and the scale and persistence of the infection. As a corollary of this type of analysis, spreading models can be used to locate influential spreaders in the network and quantify their effect on any contagion. Hu et al. [89] use the SIR model to derive a global measure of influence under a contagion, and show that this method is efficient even in large networks. SIR models have also seen use in testing the relationship between spreading processes and the formation of echo chamber structures. (e.g. [47]).

Despite the initial similarities, spreading processes for information are affected by additional factors in social interactions that require new aspects to be represented. In many cases, information takes the form of a complex contagion in which the number of times an individual is exposed to the new idea affects the likelihood of adoption. Moreover, repeated exposure can affect the duration of awareness [38]. The source of new exposure is another important factor for the spreading dynamics of information, with popular or authoritative individuals creating larger contagions [11] and contact from homophilous individuals leading to a higher rate of adoption [30]. These additional considerations facilitate new perspectives on information spreading processes, and further incorporation of such techniques in future research is likely to be fruitful.

The other main approach for analysing information spread through social systems is through agent-based modelling, where a series of autonomous actors are characterised by a series of behaviours. Uses of such methods have typically been successful, but often highlight that different information circumstances require adaptations to their modelling approach (e.g. [155]). Variants of agent-based modelling techniques consider the contagion as a series of cascades, and have again proved valu-

able for understanding the processes behind polarisation in social networks [154].

Other approaches have also been applied to understanding patterns of information diffusion. Differential equation methods help derive theoretical estimates and the interplay between degree and different timescales (e.g. [81]), and attempts to adapt the model away from uniform behaviour have shown that changing activity presents advantages for reach [9].

In the context of social media, these various methods have been used to track the spreading of hashtags, memes and discussion topics. Tracking hashtags over time has been used to measure collective attention (e.g. [111,117]) and how different topics capture different types of attention [157]. More generally, the spread of these ideas and themes can be tracked and predicted. Arif et al. [11] used several different approaches to track the spread of rumours on social media around a hostage event. They identified notable differences in how rumours spread depending on the timing and source of fact-checking efforts. Weng et al. [193] were able to predict the future virality of memes based on the characteristics of their early adopters and the structure of interaction networks. Understanding how information spreads through social media in this way is an important consideration given the potential for effective communication strategies to limit the negative consequences of polarisation in modern society.

This discussion of information spreading on social media has not yet touched on one of the most notable aspects of social media over the last few years, namely the rise of misinformation. Understanding how misinformation spreads, and how platforms and individuals can combat it, is becoming ever more important with increasing offline consequences (e.g. during the COVID-19 pandemic[2]). Many of the techniques already seen can be readily applied to understanding misinformation and the characteristics that define its success [175]. More important however is comparison of misinformation content to accurate information. Del Vicario et al. [50] found that echo chambers played a part in furthering the spread of both regular information and misinformation on Facebook, but a key difference was observed in misinformation remaining in the system for longer. This difference in tempo-

[2] `https://www.bbc.co.uk/news/stories-52731624`

ral trends between misinformation and factual content has been noticed in other contexts. Maddock et al. [121] analysed framing of rumours on Twitter over the course of several events, finding that early in the event's life cycle misinformation frames were dominant, but were subsequently challenged by fact checking efforts to address misunderstandings. Alarmingly, studies over recent years have suggested that political views can affect exposure to misinformation [77] and that the general public have become less able to distinguish misinformation [162], possibly due to bias in the diffusion of fact-checking content [169]. Attempts to tip this balance on social media sites have seen mixed success. Facebook published a list of guidelines to help identify misinformation content[3], which have been tested in experimental settings. Guess et al. [79] found that these guidelines increased discernment through increasing the scepticism of users, but such changes were ultimately short-lived in most cases.

2.2.4 Collective attention

We've seen now how scholars can use network science and social media data to understand the interplay of opinions and social structures, and how information can propagate through groups of people, but thus far we have not touched on how individuals and groups react to new events and information. We can cover this through the lens of attention and more specifically how attention behaves across larger groups.

The mechanisms used to study collective attention vary depending on the particular aspect of interest. Timeseries methods typically discard the relational information included in social media data in favour of capturing temporal trends at finer resolutions. This has seen much use in the study of specific events signifying important periods of collective attention (e.g. [111, 130, 177]). Other scholars have made use of features in user or content networks to identify structural responses to collective attention. Many consider events as multiple network snapshots to compare changes over time (e.g. [26, 74, 84]) but the challenge of determining an appropriate

[3] `https://www.facebook.com/help/188118808357379`

timescale for these snapshots makes it difficult to develop a standardised methodology. Regardless of the methodology chosen, the goal is to understand attention as revealed by the features of how people and content interact.

Over the years scholars have looked at many aspects of how attention builds around popular topics, or more globally in social systems. In a comparison of whether growth during events is focused on individuals, or across the population, Lin et al. [114] investigated how interactions were distributed through the Twitter population, and found that during the 2012 US presidential election it was more common for individuals the see the greatest benefits. Responses to external shocks have also seen study, for example terrorist attacks [84] and natural disasters [126]. The importance of external shocks ties into the role of novelty in promoting growth, as first shown by Wu and Huberman [199]. Exploring another aspect of attention, Garimella et al. [74] consider topic controversy and how it is affected by attention. Tying controversy together with polarisation, they found that both increased under collective attention. In a similar vein, Mitra et al. [130] consider the credibility of different types of information, revealing an effect with less credible information seeing more discrete attention peaks that typically reached fewer people.

As with the other features of social systems detailed in this section, numerous models have been applied to collective attention. Past work by Weng et al. [192] showed that limited attention for individuals and differing levels of attractiveness in items was sufficient to capture many trends in collective attention. This approach was subsequently expanded on by Gleeson et al. [75] to incorporate a concept of memory, with content selection occurring only from the most recently seen items. Through such an approach, a bias towards more recent information was found to produce the behaviour closest to the observed patterns. Twitter and Facebook are not the only platforms to see study in this context. Sites such as Wikipedia include promotion mechanisms that highlight certain content to visitors to their homepage, a factor included in the attention model by ten Thij et al. [177] and gives an understanding of how this affects page views. Other techniques applied to model attention can capture how it snowballs. A good example of this is the use of

Hawkes self-exciting processes by Medvedev et al. [123]. This approach incorporates a measure of attention decay, as a complement to the limited memory of Gleeson et al. [75] which does not define the length of time an item is relevant for a priori. Beyond differences in attention at the individual level, models can also reflect trends across cohorts or topics. A key aspect in this area is the initial impulse leading to an attention event. Recent work in this direction by De Domenico and Altmann [49] explicitly tied the dynamics of bursts in attention to the network formation process of preferential attachment and the analogous concept of preferential attention before determining both to be required for bursts leading to collective attention.

One area of collective attention that needs additional exploration is the temporal evolution of attention and how this can prove informative about the nature of the event. One of the earliest attempts to describe different types of collective attention events comes from Lehmann et al. [111]. They consider the patterns in daily usage of hashtags around the peak of activity, identifying four attention classes based on how hashtag usage builds up and declines. These four classes have intuitive links to types of event and potential driving forces behind attention, but the choice of daily resolution forces a timescale for the events considered. Beyond using the temporal patterns to classify types of event, temporal patterns in attention have also been used to predict events early in their lifetime. Several of these approaches use temporal networks, with changes in topology being linked to key points in the life cycle of collective attention events. Two such examples make use of community structure either to predict future hashtag use [194], or alongside the complementary structural measure of nestedness to predict upcoming peaks of attention [26]. The final temporal aspect of collective attention to touch on is the change in event patterns over longer periods spanning many events. Work in this area by Lorenz-Spreen et al. [117] examines how the trajectories of Twitter hashtag usage have changed over the years. Notably the average peak size was found to be stable, but the rate growth and decline around the peak have both increased. Moreover, comparison with other platforms show generally similar trends.

In this chapter, we have seen many different ways to study collective attention.

The tools of network science in particular have proven to be useful for understanding the social processes driving the development of collective attention. Network structure has been one of the primary mechanisms through which collective attention has been studied, and there is scope for additional work to consider the dynamics of such structures. This is something that needs to be addressed for some of the artefacts of collective attention, such as polarisation, which are not always directly linked to attention processes. The use of network and contagion modelling approaches to understand the evolution of collective attention processes presents another valuable opportunity to find new insights into this phenomenon. In the next chapter, we will use these observations to identify a series of gaps in the literature and outline how we will address them later in the thesis.

Chapter 3

Motivation

Through the methods and studies in the previous chapter, we have begun to see how networks can be used to represent different systems and how they are particularly useful for examining interactions through social media. Four key themes in collective attention emerged as of interest from the literature that we discussed in greater detail: polarisation; structural features of networks; how events and attention affect social networks; and temporal patterns in a changing world. These themes are not completely distinct, for example, events always have a temporal component, but since they can consider the same phenomena from different perspectives it is worth touching on them separately.

3.1 Emerging research themes

It should be clear from the various studies we looked at in the previous section that polarisation in social networks exists in many contexts and can be seen in many different types of interaction. One aspect of polarisation that has been relatively understudied is the impact of notable events spurring interest in the topic. Longer term analysis has been carried out (e.g. [53, 54]) but it is more common to treat the study period as a single interval with no division into smaller temporal units. While longer periods can be useful in the sense of capturing the limits of aggregate interactions, the importance of shorter intervals should not be overlooked. The

difference is likely to be particularly important for organisations that wish to bridge the polarised divide: the efficacy of short term campaigns compared to long term will have a significant impact on their strategies. Some work has begun in this direction (e.g. [27, 72]), but more is required.

One of the most common ways to measure polarisation is through measuring types of structure within the network. Typically we saw this through community detection, but methods at other network scales such as clustering [74] or centrality measures [178], or non-network methods such as text modelling [15] are also applicable. We will now be focusing on modularity-based community detection given its established use and suitability for large networks. It is worth discussing the use of modularity-based community detection in practice, particularly in the interaction with bipartite networks. As we've seen in the previous section, bipartite networks and their projections need adaptations for optimal performance of the community detection algorithms, therefore the choice of which approach is best is not always straightforward.

Many of the studies mentioned in Chapter 2 that use real-world data are framed to understand how the systems studied respond to a particular event. In some cases, this response has even been considered as a means of event detection (e.g. [26, 126]), but some distinction should be made between event detection and attention. The second study cited here uses a social sensing approach to enrich the detection of earthquakes and their impacts. This is indeed valuable work, but gives no insights into how the online social system of Twitter is responding to the event of an earthquake, instead measuring only the system outputs. In contrast, studies like that of Borge-Holthoefer et al. [26] examine the dynamics of the system in much greater detail. Their comparison of the modular and nested structures in the network over time gives an interpretation of how the change in attention affects Twitter use. Initially, the network exists in a regime of high modularity and low nestedness which suggests that users communicate in comparatively isolated groups with infrequent communication across group boundaries. As awareness of the event grows, the social network transitions to a regime of high nestedness and low modularity which

suggests that the shared topic has blurred the boundaries of the groups as they now communicate using the shared terms related to the event. Understanding the underlying network processes in this way improves the wider scientific knowledge of how such systems behave, with potential applications for managing the negative aspects of social media that have become apparent in recent years. In a similar direction, attempts to classify event types are also informative about the dynamic processes driving events. The best known of such efforts is by Lehmann et al. [111] which shows that the distribution of activity either side of the peak divides events into four well-defined classes, although this work only makes limited inferences about the process controlling such activity. This limitation is one that has persisted, and there is much scope for experimental and theoretical studies that link classes of event with dynamical processes.

The final theme that emerged in Chapter 2 was the study of networks over time. Although many studies of social media consider extended periods of time as a single snapshot of the system, there are those that consider how temporal processes are reflected in various network metrics. Such methods have been shown to allow tracking linguistic topics over time [118] or affiliations of politicians on social media [188] but are most often used as predictors of topic success (e.g. [194]) or for modelling responses to events [114]. There are however limitations in how these temporal aspects have been considered in the past. One major choice is the timescale for the analysis to be carried out on, which can have significant impacts on the types of events that can be seen. Furthermore, it is difficult to compare events at different timescales which adds significant complications to analysis of similar system features in different contexts.

3.2 Research questions

We now use these highlighted gaps in the literature to find a direction for the research undertaken in the following chapters.

The first gap in the literature we highlight ties together the themes of polarisation, temporal analysis and the impact of disruptive events on attention in social

networks. As we've seen from the work of other scholars, polarisation has been observed online for many years, particularly on the topic of climate change [54, 93]. As a topic not always at the front of the news cycle, it is also easier to identify times when climate change is receiving additional attention. Often, these situations occur around major news stories such as scientific advancements and political decisions. One such political decision was the announcement by then-President Trump of his intention to withdraw the US from the Paris Agreement on climate change on 1st June 2017. Moreover, climate change is a politically divisive topic in the US, with such an announcement gathering much awareness beyond typically interested persons. Studying this period provides an opportunity to understand how such exceptional events influence the existing attention structures in online social media. Our first research question will tackle this in the context of the sharing of information sources online.

- RQ1: How do structural patterns and polarisation in attention to information sources change under a disruptive event?

In Chapter 5, we use this politically charged period to assess how the disruptive mainstream media event affects attention to information sources and polarisation along the political and climate change dimensions. In order to measure these changes, we construct a bipartite network of Twitter users and the URLs they share, before projecting onto a network of URLs linked when they have been shared by the same user. Community structure on these projected networks is calculated for seven weeks centered on President Trump's announcement, with human-coded ratings of political and climate change bias gathered for prominent domains and used to measure the levels of polarisation. Answering RQ1 will begin to fill the gap in the literature for temporal analysis of polarisation and community structure in the shorter term as manifestations of group divisions in collective attention.

As we have seen previously, unipartite projection is one method to make the study of the richer interaction representations in bipartite networks more tractable. While projection allows for the use of more traditional unipartite network techniques for any subsequent analyses, we know that the inherent clique composition and

density of projected networks need careful consideration. In particular it's not clear which of the many projection techniques is the best approach to use when taking the unipartite projection. In order to answer RQ1, we must first address this gap in the methodology.

- RQ2: Which unipartite projection method allows community detection to best reflect ground truth in the bipartite network?

We will add to work by Li and You [113] that compared how various network metrics were affected by the choice of edge weighting scheme used in the projection process, but instead focus on the specific use case of community detection. In particular, we seek to understand how well the community structure detected on a projected network represents the community structure of the underlying bipartite network. Answering this question is important since this technique has frequently been used to study large bipartite networks given the limited availability of efficient community detection tools for bipartite networks. In Chapter 4 we explore this problem in detail. Note that we will answer RQ2 first in order to make use of the knowledge gained to support the work in Chapter 5. We begin by demonstrating the difference that the choice of edge weighting can make on detected community structure in a real-world example. Then we use a series of synthetic bipartite networks with known community structure to measure how well community detection on the unipartite projections reflect the ground truth, before using these findings to provide a recommendation of best practice for future studies applying community detection to unipartite projections.

An event of lasting impact occurred early in 2020 with the growing spread of COVID-19 and the eventual declaration of a pandemic by the World Health Organisation. As a result of this public health emergency, citizens in many countries across the world were subject to various restrictions aimed at combating the spread of infection, from social distancing measures in public and the enforcement of mask wearing through to periods of lockdown in which people should only leave home for necessities such as food or medical needs. Social media is often implicitly linked to the offline social interactions of users, but under these pandemic restrictions, many

forms of social contact are curtailed. What remains to be understood is how these offline social restrictions affect the online manifestation of personal interactions. Additionally, the varying and repeating levels of pandemic restrictions provide the circumstances to test whether repeated, similar events produce the same responses in online social networks. We consider these points in the following research questions.

- RQ3: How do restrictions in offline spaces disrupt collective attention online?
- RQ4: Do repeated, similar events generate the same collective attention effects?

Given the growth of certain online communication platforms[1] and the breadth of news stories highlighting attempts to maintain contact online[2], the level of social interaction in online settings may not reflect the same drop seen offline. Through the use of Twitter data we are able to address these questions. In Chapter 6 we make use of a dataset of geolocated tweets from England and Wales to measure how the different levels of social restrictions introduced to combat the spread of COVID-19 affect attention to locations and topics online. For this purpose, the COVID-19 pandemic is unique in recent memory through becoming the dominant news story for such an extended period of time and its relevance across all parts of society. The sustained relevance and impacts of the pandemic present a unique opportunity for scholars of collective attention online to explore how trends evolve over periods longer than the typical hours or days.

In the context of climate change and the COVID-19 pandemic we have seen two very different attention events and sought to understand their impact on the networks of social interaction online. If we wanted to compare these two events directly, we could try to characterise the event types using the approach proposed in [111]. In doing so, we face a major methodological challenge however: the response to President Trump's announcement about the Paris agreement takes place over days and weeks, whereas the COVID-19 pandemic unfolds over weeks, months, and potentially years. Given this difference in timescale, no obvious choice for bin widths

[1] `https://www.bbc.co.uk/news/business-52884782`
[2] `https://www.nytimes.com/wirecutter/blog/coronavirus-socializing-online/`

exists for characterising behaviour either side of the event peak. This motivates the following research question.

- RQ5: Can collective attention events be measured in a way that allows comparison across different scales?

Chapter 7 seeks to answer RQ5 and address this challenge by proposing a new methodology that does not require binning at temporal resolution. At the same time as removing the need to apply a temporal binning, we normalise the activity such that comparisons can be made across events attracting different levels of Twitter activity. Updating the methodology in this way provides benefits beyond just facilitating comparisons that cannot otherwise be made. As a result of the normalisation process we instead compare the arrival rates of tweets, and therefore move closer to analysing the dynamic processes driving the rates of tweets on certain topics. In addition, we will show that the choice of bin size imposes a limit on the classification of certain types of event.

3.3 Synthesis

Through a detailed examination of the literature surrounding collective attention in online social networks we have identified the following research questions:

- RQ1: How do structural patterns and polarisation in attention to information sources change under a disruptive event?
- RQ2: Which unipartite projection method allows community detection to best reflect ground truth in the bipartite network?
- RQ3: How do restrictions in offline spaces disrupt collective attention online?
- RQ4: Do repeated, similar events generate the same collective attention effects?
- RQ5: Can collective attention events be measured in a way that allows comparison across different scales?

Each of these research questions reflect themes in the understudied topic of how collective attention changes over time. Addressing this gap in the literature will promote further consideration of the inherently dynamic nature of collective attention, beyond previous attempts using static snapshots. We move towards this goal in terms of both methodological advancements and experimental studies. In the experimental contributions, we indicate two means with which the tools of network science can be used to measure the response to online social networks before, during and after disruptive events. The methodological contributions provide additional tools and advice on best practice to support future investigations into collective attention in online social networks. Ultimately, we aim to show that continuing to explore the dynamics of collective attention reveals insights that cannot be gathered from static views alone.

Chapter 4

Is it correct to project and detect? How weighting unipartite projections influences community detection

Abstract

Bipartite networks represent pairwise relationships between nodes belonging to two distinct classes. While established methods exist for analyzing unipartite networks, those for bipartite network analysis are somewhat obscure and relatively less developed. Community detection in such instances is frequently approached by first projecting the network onto a unipartite network, a method where edges between node classes are encoded as edges within one class. Here we test seven different projection schemes by assessing the performance of community detection on both: (i) a real-world dataset from social media, and (ii) an ensemble of artificial networks with prescribed community structure. A number of performance and accuracy issues become apparent from the experimental findings, especially in the case of long-tailed degree distributions. Of the methods tested, the "hyperbolic" projection scheme alleviates most of these difficulties and is thus the most robust scheme of those tested. We conclude that any interpretation of community detection algorithm performance

on projected networks must be done with care as certain network configurations require strong community preference for the bipartite structure to be reflected in the unipartite communities. Our results have implications for the analysis of detected community structure in projected unipartite networks.

4.1 Introduction

Bipartite networks are a useful representation of many real-world systems where well-defined relationships exist between two distinct classes of nodes, such as scientific papers and their authors, or digital media and the people who share it. The complexity, and relative obscurity of methods to analyze bipartite networks leads to frequent use of a unipartite projection of the system, so that more established unipartite methods can be applied (e.g. [10, 52, 135]). Another motivation for such analysis arises in situations where one class of nodes is used only to infer relationships between the other through projection (such as applying a network of users and reviews to identify groups of review spammers [184]). Unipartite projection encodes the edges between the two modes as a new network with only the nodes from a single mode, where nodes with a shared neighbor in the bipartite network are now directly connected. A cornerstone assumption is that the projected network retains key relationships such that community detection algorithms are able to capture structures which are meaningful in the bipartite context.

There are three main reasons for the use of unipartite projection in the study of bipartite networks. Firstly, methods for directly analyzing bipartite networks are limited in their scalability and their availability. Such methods, specifically designed to account for the additional complexities inherent in bipartite networks, are not widely included in popular network analysis packages and where such tools do exist they are not as capable of handling the large scale datasets of modern network science. Taking the unipartite projection allows scientists to leverage the existing toolkits for unipartite networks. It is worth mentioning that it is possible to apply unipartite methods to bipartite networks by effectively discarding their bipartite structure. In the case of community detection, this approach is less accurate than

using projection-based or bipartite methods [13], but does present an alternative means of handling large bipartite networks. In addition, direct application to bipartite networks violates the assumption of edge independence in the definition of modularity [138]. The remaining two points motivating the use of unipartite projection are best framed within use cases. In many experimental settings, one of the two bipartite modes is the primary focus. For example given a network of authors and publications we may study coauthorship using publications solely to infer edges between authors. Hence projecting the network focuses on the specific area of interest. This ties closely into the final reason to consider how unipartite projection affects community structure - some unipartite networks are implicitly projections of some hidden bipartite network. Consider again the coauthorship network. At first glance this is a unipartite network, but it is in fact an implicit projection of a bipartite network between scholars and the institutions and events they have visited - it is very unlikely for coauthorship to arise without such a meeting.

In previous work, we studied the efficacy of bipartite community detection using unipartite projections [33], constructing an ensemble of synthetic bipartite networks with imposed community structure and attempting to recover the structure from unipartite projections made with four candidate projection schemes. Here we extend our previous assessment by including three additional unipartite projection schemes. We also present a comparison of communities found using the seven projection methods applied to a real-world dataset, in this case a bipartite network linking web pages (URLs) to the Twitter users that shared them during one week of conversations about climate change.

Several community detection methods have been shown to be effective at partitioning small bipartite networks; Barber [21] adapts the null model used to compute modularity [138] on unipartite networks to the bipartite case to account for the additional requirement that the vertices incident to each edge must be in different modes. Beckett [23] reports other approaches, including weighted bipartite modularity maximization. However, optimization of bipartite modularity (e.g. through implementations such as the MODULAR package [122] or those reported by Beck-

ett [23]) is computationally demanding, and may be of limited use on large networks, such as those from online social media. Beyond algorithms using a modularity maximization approach, efforts have been made to extend stochastic block model (SBM) methods to bipartite networks. Larremore et al. [105] (and more recently Wyse et al. [200]) have demonstrated that enforcing the two node types required of a bipartite network allow the discovery of meaningful community structure. Many of these bipartite community detection methods find clusters which contain only a single node type and as such there are no edges within each community. This behavior is advantageous since it allows different numbers of communities to be found in each mode with many-to-one correspondences between them, but counter-intuitive from the perspective of unipartite communities. Despite the quality improvements achieved using bipartite community detection methods, they are less widely implemented in network analysis packages, furthering the appeal of unipartite projections.

Bipartite networks are an intuitive representation of social media activity, such as where Del Vicario et al. [52] examine how Facebook users interact with information related to the 2016 EU Referendum in the UK as two network modes. By computing the unipartite projection onto page nodes they identify communities of pages within which groups of users more frequently interact. Schmidt et al. [164] use a similar methodology to identify and explore the user groups formed around frequent likes or comments on the same Facebook content. Twitter is another platform readily studied using the unipartite projection approach. Williams et al. [197] explore behavior patterns amongst Twitter users and the news articles they share through a projection onto the article network. Analysis of this projected network found communities of news domains that were frequently shared by the same users.

Such analysis is also suited to physically-embedded networks such as where Chen et al. [40] study one representation of the Chinese bus transport network as the projection onto both modes of a stop-route network. Srivastava et al. [171] apply projection to bipartite networks of documents and terms to find clusters of similar documents. They use a threshold approach for the unipartite edges, discarding those that have a weight lower than a fixed value. Alzahrani and Horadam [10]

apply unweighted projections to two crime-related networks, finding a topographical division between urban and rural municipalities when looking at crime in New South Wales, Australia, and communities encompassing training links in a terrorist-activity network. Isah et al. [91] study the network of people and crimes to find different types of organisational structures among perpetrators. Yan and Ding [201] study various networks arising from authorship and citation behaviors and assess the similarity between topic, coauthorship and citation networks.

These are all examples where a bipartite network is explicitly projected, but numerous other studies encode this process in the network construction such as Starbird [172] who studies alternative news domains shared by Twitter users around mass shooting events. Newman [138] considers book co-purchases when testing a spectral method for community detection, which is an implicit projection of the user-item network. The design choices implicit in construction of these unipartite networks are subject to the same biases and pitfalls inherent in projection schemes.

Although unipartite projection and community detection sees frequent use in empirical work, limited theoretical study has been devoted to how the community structure in a projected network relates to the community structure in the original bipartite network. Everett and Borgatti [58] show that while unipartite projection onto a given mode results in a loss of information (since encoding one class of nodes as edges between the other is generally not reversible), it is possible to derive meaningful results by considering projections onto each mode simultaneously. Melamed [124] takes the concept of dual-projection to refine the community detection process on the bipartite networks by incorporating information from both unipartite projections. Arthur [13] extends this through a comparison of modularity metrics by including a novel modularity formulation that accounts for structure in the bipartite network. As noted by Newman [135], high-degree nodes in a bipartite network contribute a disproportionate number of edges to the corresponding unipartite projection; a node of degree k contributes of the order k^2 edges to the projection. Certainly we must be careful when weighting these edges. Guimera et al. [80] define a model to generate bipartite networks with a fixed community struc-

ture, where the parameter p denotes the fraction of network edges which join nodes within prescribed communities. They also adapt the standard definition of modularity to better reflect bipartite network structure, and test this against weighted and unweighted projections. The key finding is that in some cases both unipartite- and bipartite-modularity have similar performance. Bongiorno et al. [25] incorporate statistically validated networks into the community detection process by finding stable cores within bipartite communities. Li and You [113] examine whether any network metrics are affected by the unipartite projection process, and find that certain metrics such as clustering coefficient vary with projection scheme, while degree correlation does not.

A common first step in constructing a unipartite projection is to filter out low-weight edges such as by establishing a threshold and removing those that do not meet a specified criteria, a step which can be very helpful computationally by dramatically reducing the edge density in the projection. Sasahara [160] constructs word association networks by calculating cosine similarity between word contexts and retaining only those edges that exceed a given weighting threshold. Other methods compute edge significance relative to a null model to determine which edges to keep. Grinberg et al. [77] find networks of news sources which are visible to the same people on social media through a multiscale backbone approach. Saracco et al. [159] make use of exponential random graph models to determine statistical significance of the edges in the unipartite projection, and retain only those edges that satisfy a given significance threshold. Thresholding methods can have merit in certain use cases, but as with bipartite community detection methods they are often not included in the most widely used libraries. Another potential issue with how these methods have been used in the past is the binarizing of the remaining edge weights. While the use of thresholding is likely to increase accuracy over the case of a fully binarized projection, the loss of information in the significant edge weights is not to be overlooked. The final concern with removing edges given some thresholding criteria is fragmentation in the network. Many applications of bipartite network projection consider only the giant component, and if sufficient fragmentation occurs, it is likely

that the size of the giant component will no longer be comparable to the size of the whole network.

The methods described so far have found particularly strong purchase in the study of social systems such as online social media and scientific collaboration. Along with many natural systems, the networks in such studies often have degree distributions with a long tail, where a small number of profoundly well-connected nodes exist in a sea of low degree nodes. The popularity and success of this approach on these systems suggests these properties are beneficial to established methods, although careful consideration needs to be given to how the properties of the bipartite degree distribution influence the structure of the projected unipartite network and the performance of community detection methods.

In this study, we consider a range of network and projection types under community detection, and evaluate the quality of the output against prescribed communities. Networks are differentiated by their degree distributions, selected to include those characterised by geometric-, binomial- and zeta-like tails, and the spread of edges within, and between prescribed communities. Unipartite projections are taken using seven different edge weighting schemes, before testing the ability of unipartite community detection algorithms to recover bipartite community structure. We first illustrate the different outcomes associated with each projection scheme applied to a real-world bipartite network. We next perform a more rigorous test that seeks to recover bipartite community structure after unipartite projection of a series of synthetic networks with imposed community preference. In Section 4.2 we detail the real-world dataset studied, the network model used to sample synthetic networks, the seven projection weighting schemes used, and the metrics by which we measure community detection performance. Section 4.3 presents the results of our experiments and Section 4.4 discusses their consequences.

4.2 Methods

This section begins by explaining the methodology to construct and study a network from a Twitter dataset. Also in this section, we outline a model for generating ran-

dom bipartite networks with a prescribed community structure and distinct degree distributions. Seven different weighting schemes are described for use in unipartite projections and finally we outline the process for computing and evaluating the accuracy of community detection on these network projections. We use *left* and *right* as generic labels for our bipartite modes throughout this paper. The results from these methods are presented in Section 4.3.

4.2.1 Real-world dataset

For our study of how the different projection weighting schemes affect real-world networks, we make use of a Twitter dataset to construct a bipartite network between users and the URLs they share. The dataset was gathered from the Twitter Streaming API[1] using the search terms *climate change* and *global warming* for a one week period between 31-05 and 06-06-2017. We keep only those tweets containing URLs which can be resolved and remove any user that shared URLs more than 50 times during the week as a likely automated account such as news aggregators. Finally, we apply a disambiguation step to URLs by following any permanent redirects to reveal the destination of masked URLs, such as those from link shortening services. This leaves a dataset of $187,378$ tweets by $54,347$ users sharing $20,880$ distinct URLs.

From this dataset, we construct a bipartite network linking user nodes to URL nodes, adding an edge whenever the user includes the URL in one of their tweets. Edge weights are assigned as the number of times a user shared the same URL. This gives a bipartite network with $80,009$ edges (numbers of user and URL nodes as above). We restrict to the giant component for all further analyses, which contains $7,496$ URL nodes, $42,113$ user nodes and $63,755$ edges. Taking the projection onto the URL nodes of the giant component gives a unipartite network with $53,652$ edges.

[1] `https://developer.twitter.com/en/docs.html`

4.2.2 Synthetic bipartite networks

Generative bipartite network model

Many different methods can be developed for producing synthetic bipartite networks; free variation of network statistics such as edge density, degree distribution, and vertex correlation may result in a wide range of structures and behaviors. Our model was designed to minimize the number of assumptions made by prescribing only the degree distribution and the vertex correlations required to impose community structure on the network. Even in this case of limited assumptions, there are many candidate degree distributions to choose from. Here we construct a generative model motivated by a simple physical interpretation, described in detail by Weaver [186].

If we consider a growth process, we add new vertices at a constant rate, with a number of incident edges governed by preferential attachment. A parameter m determines the level of preference for high degree nodes in the assignment of new edges. We study two cases: $m \to \infty$ (no preferential attachment) and finite m (a strong preference for high degree vertices). Without preferential attachment, the growth process still produces an interesting degree distribution. The main difference between the two cases is that preferential attachment yields a zeta distribution (or colloquially a "power law") as vertex degree k increases in the tail of the distribution. Preferential attachment therefore leads to the formation of extremely well-connected vertices, a feature typical of self-organizing structures in nature and human-society [134]. Real-world phenomena, such as the distribution of page interactions on Facebook [52], are well represented by this model. With no preferential attachment a geometric degree distribution emerges. These two cases for m enable a comparison of how heavy-tailed distributions of different types interact with the projection process.

Previous work by Weaver [186] derived the steady state degree distribution of randomly grown networks under different preferential attachment conditions as:

$$\mathrm{pmf}(k) = \frac{m+\delta}{m(\delta+1)+1} \frac{(m)_{\left(\frac{m}{\delta}+2\right)}}{(m+k)_{\left(\frac{m}{\delta}+2\right)}} \tag{4.1}$$

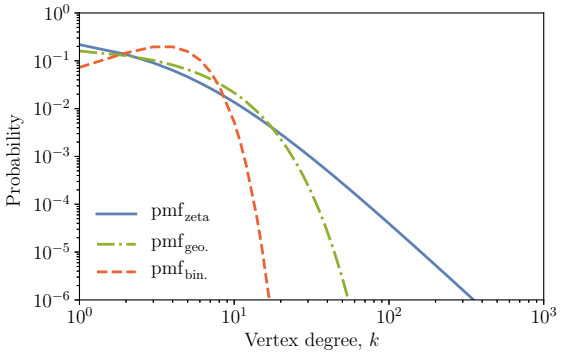
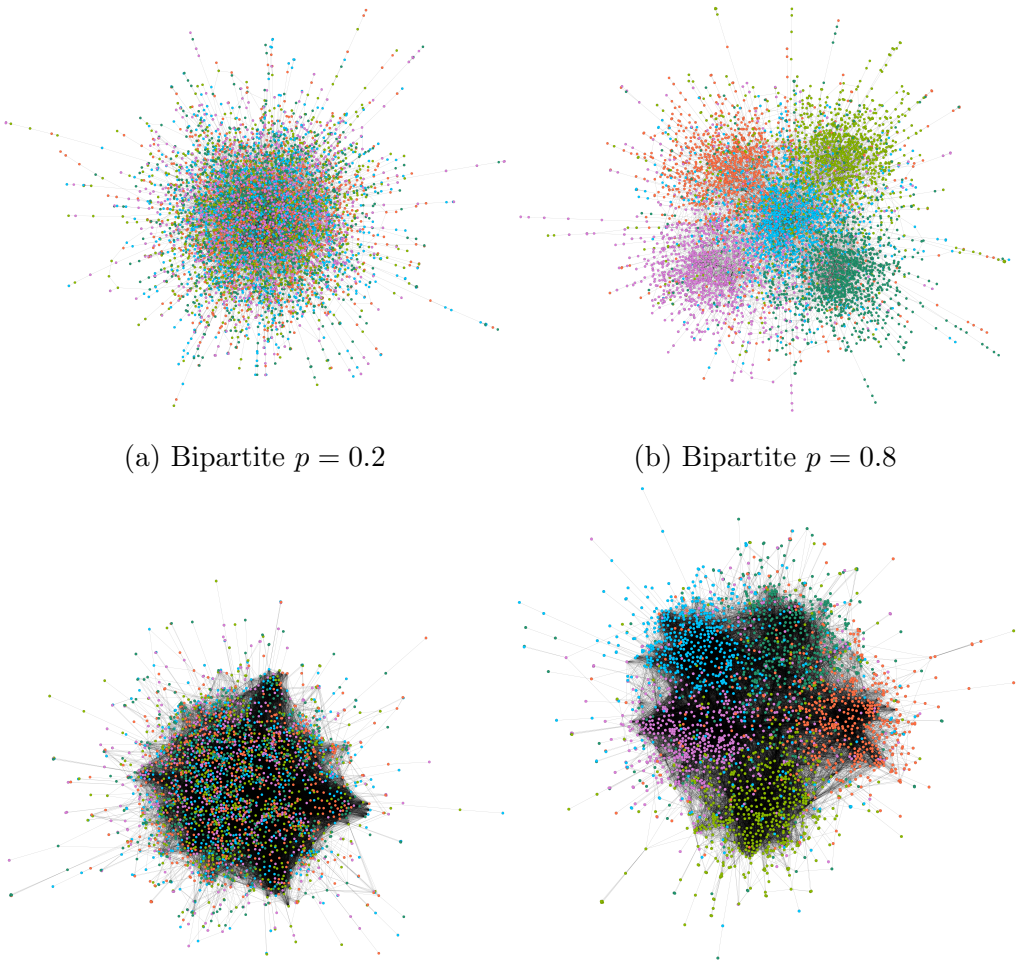


Figure 4.1: The three degree distributions used in the generative model. Each distribution has identical mean $\langle k \rangle = 4$, but vary in the weight of the tail as k increases.

making use of the Pochhammer notation $(a)_b = a(a+1)...(a+b-1)$. The parameter δ defines the number of edges incident to each newly added vertex. In the case $\delta = m = 4$ this simplifies (for sufficiently large k) to $\mathrm{pmf}_{zeta}(k) \sim k^{-3}$, that is, a zeta distribution representing preferential attachment. Setting $m \to \infty$ and $\delta = 4$ gives $\mathrm{pmf}_{geo}(k) = \frac{1}{5}\left(\frac{4}{5}\right)^k$, that is, a geometric distribution from growth without preferential attachment. To contrast these physically-motivated models, we provide an Erdös-Rényi bipartite graph for comparison, with degree distribution given by $\mathrm{pmf}_{bin}(k) = \binom{\delta N}{k} N^{-k} \left(1 - \frac{1}{N}\right)^{\delta N - k}$. Fig. 4.1 plots each of these degree distributions.

After choosing a degree distribution, construction of a random network begins by assigning the number of vertices in the left- and right-modes, $N_l = N_r = N$ respectively, and sampling degrees from the distribution for each vertex. In all cases, both bipartite modes sample from the same degree distribution. Edge creation is performed by randomly selecting pairs of nodes, choosing one from each mode weighted by their unassigned degree. We impose community structure following the method of Guimera et al. [80] by defining a partition of the vertices into M equally-sized communities before assigning edges, with a one-to-one correspondence between the communities in each mode. We define a parameter p to fix the probability of an edge connecting two vertices in the same community, with complementary probability $1 - p$ of connecting vertices regardless of their assigned communities. Notably, the proportion of edges joining vertices in the same community is not simply p, but $p + \frac{(1-p)}{M}$, which varies from $\frac{1}{M} \to 1$ as p varies over $0 \to 1$. Many vertices have degree $k = 0$, a characteristic frequently mirrored in real-world citation networks [104]. We discard isolated nodes before continuing our analysis. Sample network giant components produced by this model, and their projections, can be



(a) Bipartite $p = 0.2$ (b) Bipartite $p = 0.8$

(c) Unipartite $p = 0.2$ (d) Unipartite $p = 0.8$

Figure 4.2: The giant components of bipartite networks produced by a single network instance of 10^4 nodes with average degree 2 divided into $M = 5$ communities along with their unipartite projections: (a), (c) have low community preference $p = 0.2$, while (b), (d) use high community preference, $p = 0.8$. Node colors indicate community assignment. The dramatic increase in edge density across the projection process can be seen by the relative intensity of black edges in (c), (d). Networks are visualized using Gephi with layout determined by the ForceAtlas2 algorithm [92].

seen in Fig. 4.2.

Producing an ensemble of synthetic networks

In order to study the expected behaviour of different projection weighting schemes under community detection, we construct an ensemble of synthetic networks with known community structure as outlined in Section 4.2.2. Each network consisted of $N = 10^6$ nodes (in each mode) divided into $M = 5$ communities. For each degree distribution we fix the expected node degree to be 4, giving approximately 4×10^6 edges in each synthetic network. For each value of the community preference parameter

$p \in \{0, 0.1, 0.2, \ldots, 1\}$, networks are generated for each degree distribution outlined by Fig. 4.1. We present our results as averages over 100 network realisations for each combination of projection weighting, degree distribution and community preference p.

4.2.3 Unipartite projection and edge-weighting schemes

Taking the unipartite projection produces an edge between each node pair of the chosen class with at least one shared neighbor in the other class. A key part of this process is how edge weights are calculated in the projection; approaches can be as simple as recording presence of a mutual neighbor or as complex as nonlinear weighting from the overlap of the neighborhood sets. Each of the weightings outlined below have been designed to work with both weighted and unweighted bipartite networks. In practice however, the bipartite networks tested (both empirical and synthetic) have primarily binary edges, that is the proportion of edges with weights larger than one is small. Here we detail seven weighting schemes for edges in the unipartite network which will be tested against the quality of unipartite community detection relative to different levels of bipartite network structure. As outlined here, the methods describe projection onto the right nodes, but apply equally to the left nodes under suitable transposition. Given this flexibility in application, the choice of which projection to use is circumstantial and depends on the research question. In our case the choice is arbitrary; the generative model produces networks with statistically symmetric modes and hence statistically symmetric projections. In an experimental setting, it is often clear which of the two modes is of interest, making it obvious which nodes should be projected onto.

The *simple* weighting scheme calculates the weight w_{ij} for edge e_{ij} in the projected network as the number of neighbors nodes i and j share in an unweighted bipartite network. In the case of a weighted network, w_{ij} represents the sum of the product of edge weights on all i, j-paths of length two. Under simple weighting, the bipartite adjacency matrix B (with rows corresponding to the left-mode, columns corresponding to the right-mode and entries corresponding to the edge weights) is

used to define the unipartite adjacency matrix:

$$\mathrm{U}_{\text{simple}} = \mathrm{B}^{\top}\mathrm{B}. \tag{4.2}$$

Note that U is symmetric, and has non-zero diagonal elements and as such encodes a network with self-connections. In this paper, we do not allow self-connections, and set the diagonal elements to zero.

The *binary* weighting scheme is calculated from the simple edge weights by truncating at 1; that is we do not consider the number of shared neighbors between a pair of nodes. As such, we record presence or absence of a shared neighbor as a 1 or 0 respectively, giving:

$$\mathrm{U}_{\text{binary}}[i,j] = \begin{cases} 1 & \text{if } \mathrm{U}_{\text{simple}}[i,j] \neq 0, \\ 0 & \text{if } \mathrm{U}_{\text{simple}}[i,j] = 0. \end{cases} \tag{4.3}$$

The *hyperbolic* weighting scheme, introduced by [135], is a means to limit the influence of high-degree nodes in the bipartite network in the projected network. A node of degree k in the bipartite network will contribute a total edge weight proportional to the square, $\frac{1}{2}k(k-1)$ under the simple weighting scheme. As a result, high-k nodes can have a disproportionate influence on total edge weight and consequently community quality in the projected network. This is of particular concern in networks with long-tailed degree distributions. The hyperbolic scheme applies a scaling factor of $(k_i - 1)^{-1}$ to each edge created in the projection of node i with degree k_i. In this scheme, high degree nodes still have an increasing contribution to the total edge weight, but now contribute linearly by degree, $\frac{1}{2}k$. Under hyperbolic weighting, the unipartite adjacency matrix is defined as:

$$\mathrm{U}_{\text{hyper.}} = \mathrm{B}^{\top}\mathrm{W}\mathrm{B} \text{ where } w_{ij} = \begin{cases} (k_i - 1)^{-1} & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases} \tag{4.4}$$

The *unary* weighting scheme extends Newman's hyperbolic weighting to normalize each node's contribution to the total edge weight in the projected network. The edge weights formed by projection of node i are rescaled by $2k_i(k_i - 1)^{-1}$. As

a result, the total edge weight contribution of a node to the projected network is exactly 1. Under unary weighting, the unipartite adjacency matrix is defined as:

$$\mathrm{U}_{\text{unary}} = \mathrm{B}^{\top}\mathrm{V}\mathrm{B} \text{ where } v_{ij} = \begin{cases} 2(k_i(k_i-1))^{-1} & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases} \tag{4.5}$$

The *random walk* weighting scheme evaluates the probability distribution of two-step random walks on the bipartite networks to determine the projected edge weights. Such an approach is often used when calculating node ranking or similarity for recommender systems (e.g. [110]). This is calculated by row-normalizing the bipartite adjacency matrix and performing a matrix multiplication as with previous methods, that is:

$$\mathrm{U}_{\text{randw}} = |\mathrm{B}^{\top}|_{\text{L1}}|\mathrm{B}|_{\text{L1}} \tag{4.6}$$

where $|\mathrm{B}|_{\text{L1}}$ denotes the matrix B after L1 normalization of the rows.

The *cosine* weighting scheme is a nonlinear measure of similarity between node neighborhoods. We define the weight of an edge between two nodes in the unipartite network as the cosine similarity of the two corresponding neighborhoods, that is:

$$\mathrm{U}_{\text{cosine}}[i,j] = \frac{\mathrm{B}[:,i] \cdot \mathrm{B}[:,j]}{|\mathrm{B}[:,i]||\mathrm{B}[:,j]|}. \tag{4.7}$$

The *Jaccard* weighting scheme measures the overlap between nodes' neighborhoods. Bipartite edge weights can be incorporated by weighting neighborhood elements. The edge weight is defined as the ratio between the sizes of the intersection and the union of the node neighborhoods, that is:

$$\mathrm{U}_{\text{Jaccard}}[i,j] = \frac{|N(i) \cap N(j)|}{|N(i) \cup N(j)|}, \tag{4.8}$$

where $N(i)$ is the immediate neighborhood of node i.

We illustrate the effects of each of these weighting schemes on community detection in the unipartite projections in Fig. 4.3 by constructing a small, unweighted bipartite network such that the community structure found on each projection is

unique.

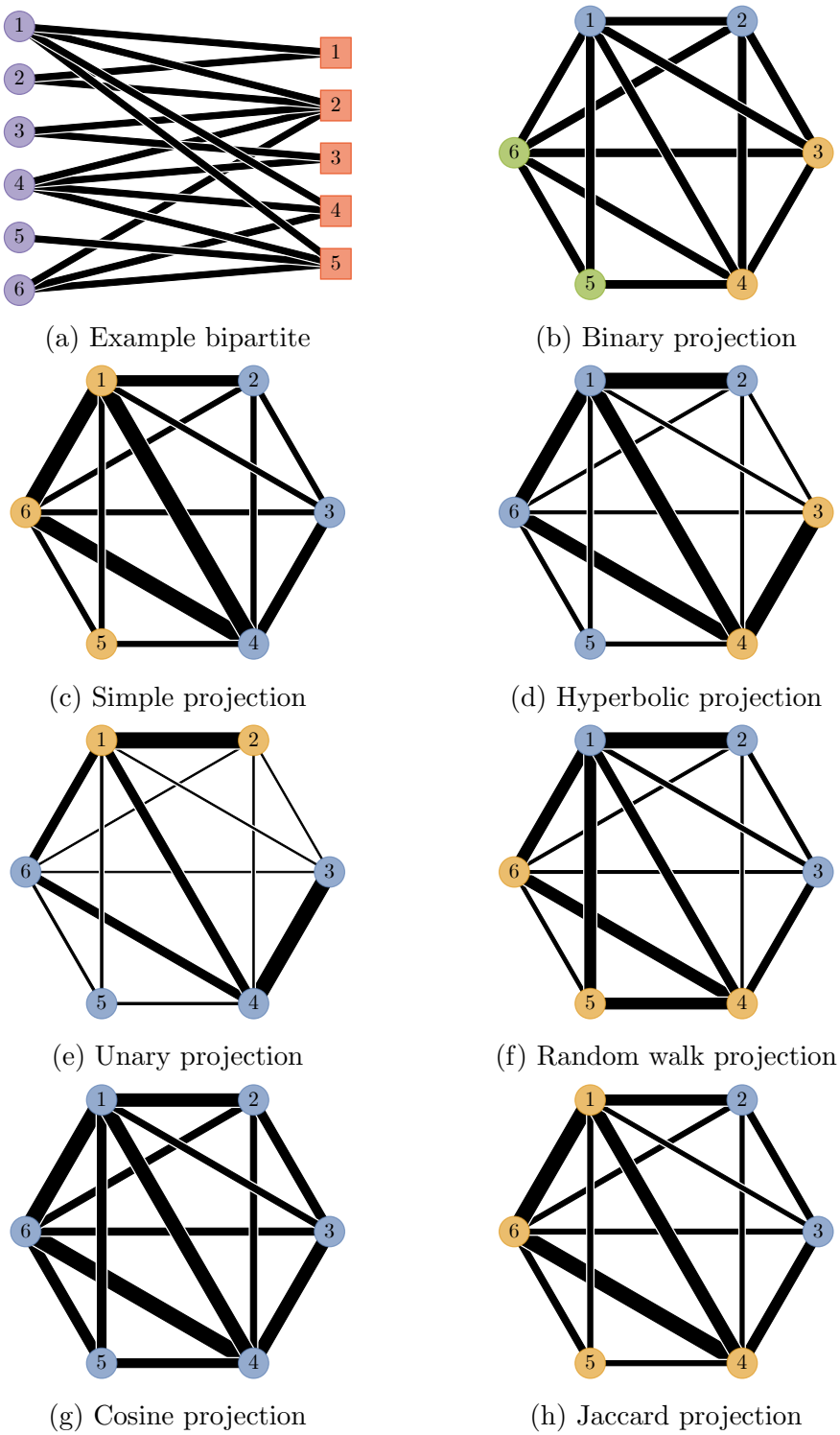
4.2.4 Community detection

We produce bipartite networks by varying the parameter p, which controls the preference of nodes to connect with other nodes within their prescribed communities. Given a unipartite projection we apply community detection with the expectation that we can recover some amount of the community structure used in construction. When analyzing the networks produced by the generative model we measure the accuracy in the returned community partition with respect to the prescribed communities.

In all cases we use the Louvain algorithm proposed by Blondel et al. [24] for community detection. This algorithm estimates the best community partition through modularity maximization on the large and locally dense networks produced by our model. The algorithm begins by assigning each node to its own community, iteratively changing the community label of a node to that which produces the largest increase in modularity among the community labels of its neighbors, with ties broken by random selection. When no more steps can increase modularity, a new network is induced by merging all nodes in a community into a single node and the first step is repeated on the new induced graph. This method proves highly scalable, allowing calculation of communities in large, weighted networks. In our case, we consider unipartite modularity, but by changing the modularity function the algorithm can be applied to other network types. The Louvain algorithm requires no information about the number of communities to find. This behaviour is ideal for many experimental use cases as it means the final partition is decided entirely by network topology.

4.2.5 Assessing the accuracy of community detection

Our detected network communities are evaluated against the prescribed community labels by using the adjusted Rand index, that is the proportion of all node pairs for which community labels are either the same or different in *both* the computed and



(a) Example bipartite

(b) Binary projection

(c) Simple projection

(d) Hyperbolic projection

(e) Unary projection

(f) Random walk projection

(g) Cosine projection

(h) Jaccard projection

Figure 4.3: A small, unweighted bipartite network and its unipartite projections. The node color in the bipartite denotes the left and right modes, the node colors in the unipartite projections denote the communities found by the Louvain algorithm [24] and line thickness is proportional to edge weight. The bipartite network was constructed to ensure the detected community structure is different under each projection to highlight the impact of projection weighting on community detection.

prescribed labelling, adjusted by the expected level of agreement by chance:

$$\mathrm{ARI} = \frac{\mathrm{RI} - \mathbb{E}(\mathrm{RI})}{\max(\mathrm{RI}) - \mathbb{E}(\mathrm{RI})}. \tag{4.9}$$

The adjusted Rand index takes values in $[0, 1]$, where values close to 0 indicate that agreement between the true and detected communities is no better than chance, and 1 indicates that the true and detected communities are identical. This measure has the desirable property that the precise community labels found are not important for evaluation of the adjusted Rand index, only whether two given nodes have the same community assignment in the detected and reference structures. As a result, permutation of the community labels does not affect this measure. We choose the adjusted Rand index over other information theoretic measures for two reasons. We argue that it is important for any comparison of our community structures to account for chance agreement. The adjusted Rand index explicitly accounts for this using a null model, whereas competing measures such as normalized mutual information do not. We also follow the advice of Romano et al. [156] who find that the adjusted Rand index performs better when considering relatively few different labels in the reference partition.

We also compare the sizes of the detected communities to the sizes known in our synthetic networks. We do this by computing the expected community size of a random node, that is

$$\frac{\sum_C |C|^2}{(\sum_C |C|)^2} \tag{4.10}$$

where $|C|$ denotes the number of nodes in community C. Note that this measure is a rescaling of the mean community size which can be heavily skewed by variable sizes of the giant component over different realisations of the random network model.

4.3 Results

4.3.1 Example: Real-world network from Twitter

We first report on the communities found in unipartite projections of the Twitter network made using the different weighting schemes. Fig. 4.4 visualizes the communities found in the seven projections of the giant component, with a layout determined by the ForceAtlas2 algorithm [92]. Node colors depict community membership, sorted by community size. Three types of community assignments appear across the seven projection weighting schemes. In the first type, the largest community dominates the left-hand cluster of the network (as seen in Fig. 4.4a and Fig. 4.4g). In the second type, the largest community dominates the right-hand cluster (as seen in Figs. 4.4c, 4.4d and 4.4f). In the final type, each cluster is made up of multiple smaller communities.

As we do not know the "true" community structure for the empirical Twitter network, we cannot compute the accuracy of community detection. Instead we can compare the community partitions found by different projection schemes with each other. Table 4.1 shows the expected community size, the sizes (number of nodes) of the 5 largest communities detected and the corresponding modularity. The random walk, cosine and Jaccard methods stand out as producing extremely high modularity scores along with a large number of small communities. This seems counter intuitive and as we will see in Section 4.3.2, a high modularity in a projected network is not always a sign of underlying community structure. The binary weighting finds the single largest community, but leads to a large number of smaller communities compared to the simple, hyperbolic and unary weightings which perform similarly, with a broader distribution of community sizes. Calculation of the Gini coefficient on the distributions of community sizes over the different projection weightings finds that most of the detected partitions have high size inequality ($0.75 < G < 0.9$). The exception to this is the random walk weighting scheme which had a Gini coefficient of 0.409, indicating a very broad distribution of small communities.

Table 4.2 shows the pairwise adjusted Rand index between the community struc-

Weighting	Expected community size	Size of 5 largest communities					Modularity
Binary	0.122	2124	1081	525	421	375	0.565
Simple	0.105	1377	1069	1046	834	720	0.614
Hyperbolic	0.110	1808	1156	781	530	527	0.580
Unary	0.087	1664	899	632	494	420	0.609
Random walk	0.032	829	451	333	248	238	0.791
Cosine	0.042	879	833	488	402	231	0.870
Jaccard	0.068	1631	890	164	157	153	0.932

Table 4.1: Statistics for the communities found by the Louvain algorithm [24] on the unipartite projection of the Twitter dataset under different weighting schemes.

Weighting	Binary	Simple	Hyperbolic	Unary	Random walk	Cosine	Jaccard
Binary	1	0.401	0.285	0.23	0.191	0.279	0.416
Simple	0.401	1	0.368	0.292	0.21	0.26	0.23
Hyperbolic	0.285	0.368	1	0.621	0.205	0.227	0.197
Unary	0.23	0.292	0.621	1	0.207	0.214	0.173
Random walk	0.191	0.21	0.205	0.207	1	0.327	0.196
Cosine	0.279	0.26	0.227	0.214	0.327	1	0.408
Jaccard	0.416	0.23	0.197	0.173	0.196	0.408	1

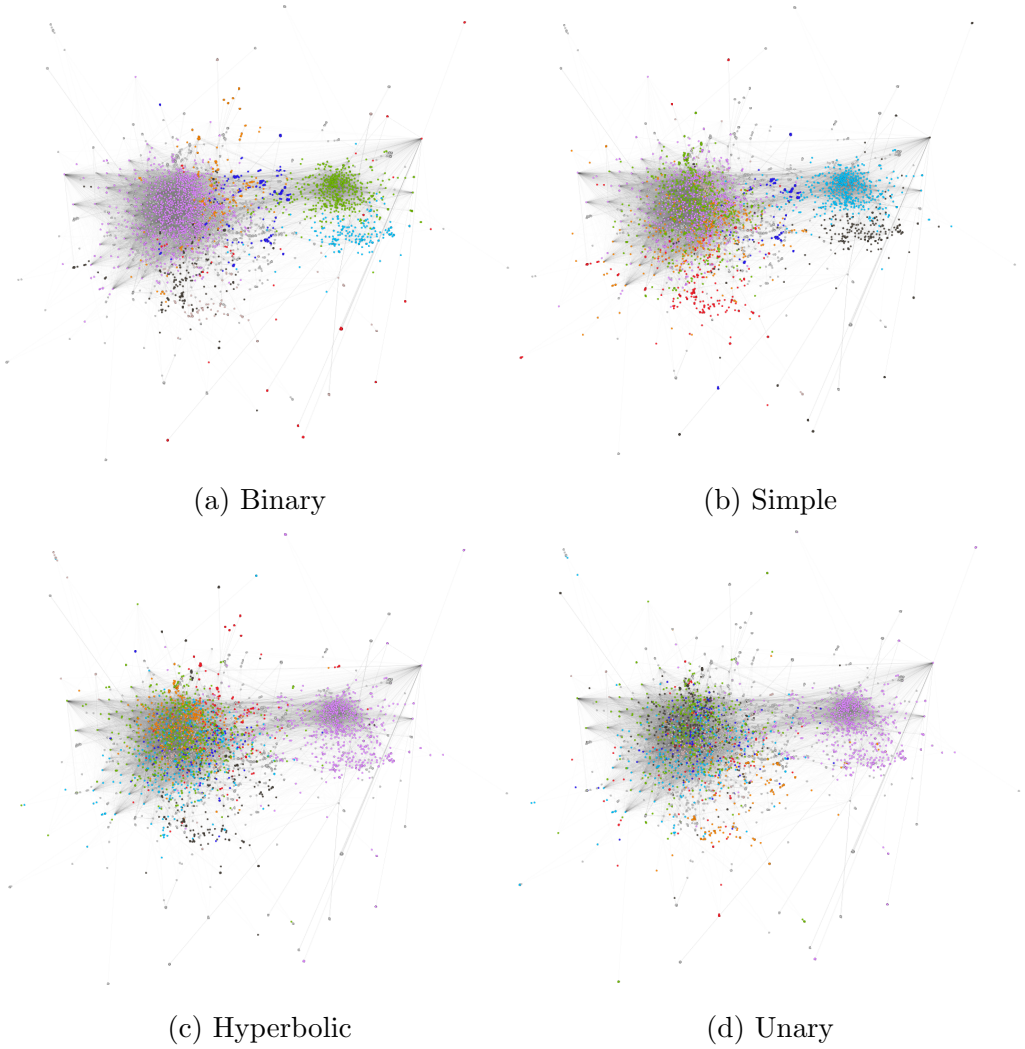
Table 4.2: Pairwise adjusted Rand index comparisons of the different community structures detected on projections of the Twitter network under the seven weighting schemes. Cell shading increases linearly with similarity.

tures detected under each projection scheme. The general trend between the community assignments for nodes shows limited similarity under the different weighting schemes. A notable exception is the hyperbolic and unary schemes which are the most similar pair. Some similarity is also observed between the binary and simple weightings and the cosine and Jaccard weightings. Also of note is the random walk weighting, which gives the lowest average similarity to other methods.

Taken together, the results from applying different unipartite projection schemes to a real-world bipartite network give a good indication that the method of projection has a large impact on the community partition that is found. We do not know the true partition of this empirical network, so we cannot determine the accuracy of community detection by each method. However, the variations between outcomes for different methods raise the question of which projection method permits the most accurate identification of community structure.

4.3.2 Testing with synthetic network ensembles

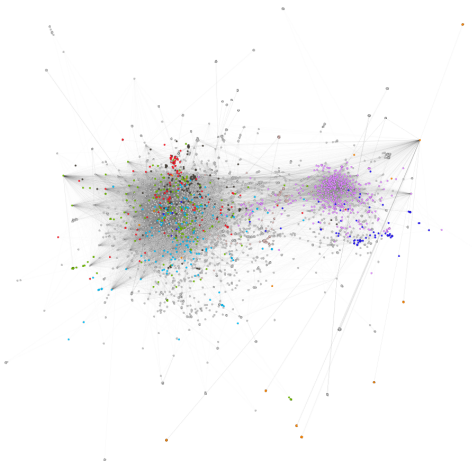
In this section, we report results from a systematic exploration of community detection accuracy using unipartite projections of bipartite networks with known com-



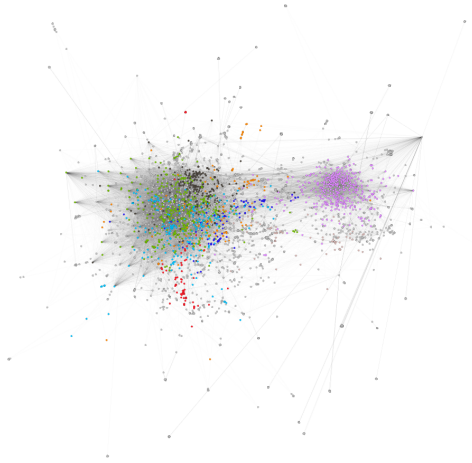
(a) Binary (b) Simple

(c) Hyperbolic (d) Unary

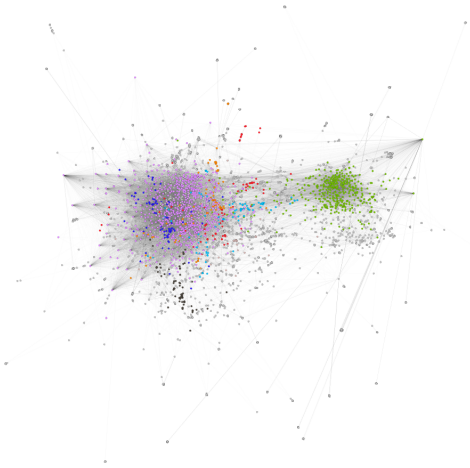
Figure 4.4: Community structure found when applying the seven different weighting schemes to the projection of the Twitter network. Only nodes of degree at least 5 are visible and node color corresponds to communities in decreasing size order (pink, green, light blue, black, orange, red, blue, grays respectively). Note the variability of the division and size ranking of different communities under each of the seven different weighting schemes.



(e) Random walk



(f) Cosine



(g) Jaccard

Figure 4.4: Community structure found when applying the seven different weighting schemes to the projection of the Twitter network. Only nodes of degree at least 5 are visible and node color corresponds to communities in decreasing size order (pink, green, light blue, black, orange, red, blue, grays respectively). Note the variability of the division and size ranking of different communities under each of the seven different weighting schemes.

munity structure. Since the two modes in our synthetic networks are generated and connected using the same processes, the left and right projections are statistically indistinguishable. As such, we report only results on the right projection. The Louvain community detection algorithm is applied to each projected network, and we evaluate the accuracy of the resulting partition using the adjusted Rand index, modularity and expected community size.

A key consideration when constructing networks from real-world datasets is sampling; how complete must a sample from one set of nodes be in order to recover a dataset representative of the system as a whole? To give a more concrete example, suppose we wish to sample the authorship network by taking all works by a number of scholars. How many authors are we required to sample to produce a network that has a giant component of the necessary size? By constructing models of bipartite networks with different degree distributions we can provide insights for a breadth of relevant networks. We iteratively sample nodes from one mode, in this case right, computing what proportion of nodes in the left mode we discover, and furthermore what proportion are connected to the largest network component. The results displayed in Fig. 4.5 show that for a geometric-tailed degree distribution, and more so for the binomial distribution, we see a similar effect to Callaway et al. [32], where a finite sample is required to produce a sample network with a giant component. In contrast, networks comprised of long-tailed degree distributions have a giant component which can be recovered from very small vertex samples as a consequence of the high degree "hub" nodes; a phenomena which may be credited in part for the success and growth of this field.

As previously mentioned in establishing the network size, an additional computational challenge exists around projecting networks with heavy-tailed degree distributions. Recall that a bipartite vertex with degree k produces $\frac{1}{2}k(k-1)$ unipartite edges after projection. This typically leads to a dramatic increase in the edge density of the projected network. With 10^6 nodes in each mode and 4×10^6 edges, the projected network from binomial, geometric and zeta distributions result in roughly 8×10^6, 16×10^6, and 120×10^6 edges respectively; that is to say the long-tailed

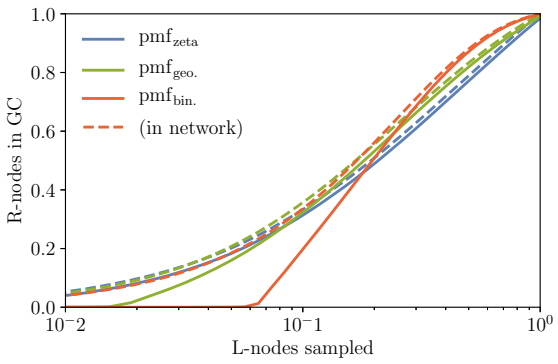


Figure 4.5: Fraction of all R-nodes included in the sampled network and its giant component for a given sample of L-nodes. These results are derived from network instances with 10^6 vertices in each of the L- and R-modes, 4×10^6 edges, and community preference $p = 0.5$.

degree distribution experiences a fifteen-fold increase in network density over the binomial degree distribution. The number of edges in a unipartite projection is tied to the second moment of the bipartite degree distribution; long-tailed distributions frequently have divergent second moments (as in this case) which cause the number of edges to grow rapidly with the size of the network, and produce dense unipartite projections.

Fig. 4.6 reports three different metrics for the performance of community detection on the unipartite projections: adjusted Rand index, unipartite modularity, and expected community size for a uniformly chosen node. The extent to which prescribed communities can be recovered computationally is strongly dependant on all of our model parameters; the node degree distribution, the strength of imposed community structure, and projection weighting. In particular, agreement with prescribed community labels is only found with strong imposed community structure at high p.

The adjusted Rand index results in Fig. 4.6a show the extent of agreement between the prescribed and detected community labels. A near-zero value indicates that community labels are a no-better predictor of true values than a random assignment while increasing values indicate better performance. Fig. 4.6a shows that for a long-tailed degree distribution, community detection reveals meaningful community labels at a much lower threshold of community preference, approximately $p \geq 0.4$ compared to values of roughly 0.6 and 0.7 for geometric and binomial degree distributions respectively. In all cases, the weighting scheme has a significant impact

on performance, with hyperbolic weighting outperforming other schemes, and unary weighting performing much worse.

The modularity (Fig. 4.6b) and expected community size (Fig. 4.6c) results provide additional insight into the recovery of the underlying community preference by community detection on the unipartite projection. The cause appears to be that modularity in the unipartite projection is not completely determined by the level of imposed community structure, as the expected increasing behavior only occurs with large p. Furthermore, modularity is shown to be non-zero for partitions on networks with weak or no imposed community structure (as high as 0.5). This suggests high quality partitions have been identified in the projected network, although in the generative method we have imposed weak or no bias at all. The hyperbolic weighting scheme demonstrates a desirable effect in this regard, returning the lowest modularity for low p, whereas the unary weighting scheme performs worst, giving high modularity for low p. In the hyperbolic and unary cases we find that the expected size of community is small when the adjusted Rand index indicates poor recovery of the imposed community structure, before converging to 0.2, the value of the true partition. Exceptions are the binary and simple weightings which decrease to the convergent value. Combined with the modularity results, this suggests that at low p the modularity maximizing algorithms find high modularity by partitioning the network into one or more large communities. More meaningful communities emerge with increasing p. This behavior is likely caused by the dominance of large cliques formed in the projection of high degree nodes. The binary and simple weighting schemes have no means of countering the impact of hub nodes when detecting communities, therefore they are often formed by the composition of multiple cliques. We also observe in Fig. 4.6 that the binary and simple weighting schemes demonstrate the most variance across the 100 iterations suggesting a susceptibility to recording different results from different observations of the same process. Across each projection method we find uniformly high Gini coefficient among the distribution of detected community sizes (> 0.75 for all p, weights and degree distributions). This shows that the range of community sizes is large, an unsurprising result given the

agglomerative nature of the Louvain algorithm.

Fig. 4.7 compares the hyperbolic, random walk, cosine and Jaccard projections. We find that the four weighting schemes perform similarly in the case of geometric and binomial degree distributions for both the adjusted Rand index and mean community size. Modularity is similar among the random walk, cosine and Jaccard projections, but is still higher than that of the hyperbolic projection at low p. Considering the zeta degree distribution differentiates the projections schemes more clearly. Fig. 4.7b shows that the cosine and Jaccard weightings perform poorly by reporting very high modularity when there is weak underlying structure in the network and showing little change as p increases. The random walk weighting reports some change in modularity as p increases. We also see that the hyperbolic weighting recovers the most information about the true network structure in Fig. 4.7a. In Fig. 4.7c, the Jaccard projection is unique in returning a structure with consistent expected community size across p, and remains close to the value of the true partition. Fig. 4.7 also shows that the cosine and Jaccard weighting schemes experience large variance over model observations, much like the binary and simple schemes. As with the first four projection methods we find that the Gini coefficient for the random walk, cosine and Jaccard weightings is uniformly high (> 0.8 for all p, weights and degree distributions).

4.4 Discussion

Our exploration of the different community structures detected under the seven weighting schemes on the real-world dataset illustrates the huge impact that edge weighting has on community detection. As in most experimental cases, the true community structure is not known for the sharing of URLs on Twitter, so we cannot assess which is closest to some ground truth. Numerous previous studies strongly suggest there is utility in this approach, so we are left in a position to decide which properties are desirable for further analysis, and assess the community quality by measuring coherence of some node properties.

The community structures reported on the Twitter dataset in Section 4.3.1

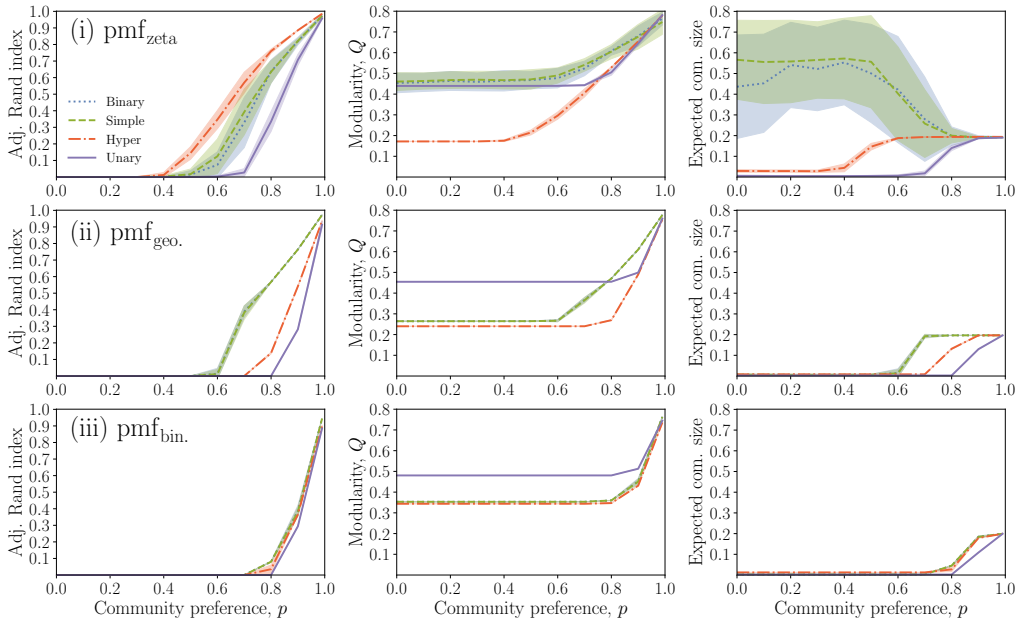


Figure 4.6: Comparison of community detection after binary, simple, hyperbolic and unary weighted projections. Lines show the mean over 100 iterations and the shaded region indicates ± one standard deviation. Left to right: (a) Adjusted Rand index, (b) modularity, and (c) expected community size, across bipartite networks with varying levels of community structure. Top to bottom: (i) pmf$_{\text{zeta}}$, (ii) pmf$_{\text{geo.}}$ and (iii) pmf$_{\text{bin.}}$ bipartite degree distributions.

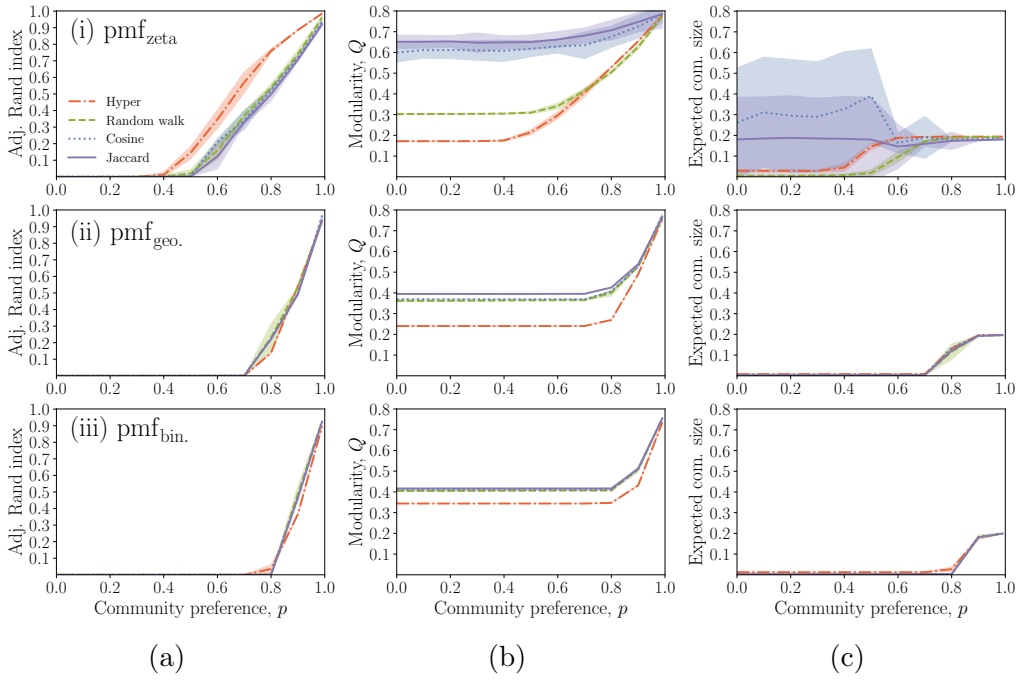


Figure 4.7: Comparison of community detection after hyperbolic, random walk, cosine and Jaccard weighted projections. Lines show the mean over 100 iterations and the shaded region indicates ± one standard deviation. Left to right: (a) Adjusted Rand index, (b) modularity, and (c) expected community size, across bipartite networks with varying levels of community structure. Top to bottom: (i) pmf$_{\text{zeta}}$, (ii) pmf$_{\text{geo.}}$ and (iii) pmf$_{\text{bin.}}$ bipartite degree distributions. The hyperbolic weighting is included here for comparison with Fig. 4.6.

demonstrate the influence weighting schemes have on the resulting community partition. The Jaccard and binary weighting schemes stand out as performing poorly, resulting in one dominant community and a large number of small communities, certainly more than can be justified by a topical or demographic argument. If such a granularity is required for analysis, it is recommended that one of the other methods is applied alongside recursive community detection, that is subsequent use of community detection on the community subgraphs.

The cosine and random walk schemes perform differently from the other weightings, finding qualitatively different community structures. Neither finds any large communities, and the cosine weighting finds many more communities than other methods. This lack of similarity with the other methods suggests that the cosine and random walk weighting schemes encode different, less intuitive network properties than other projection schemes.

When applied to the Twitter dataset, the simple, hyperbolic and unary weighting schemes perform similarly, finding similar communities both in size distribution and labelling. Under the unary weighting sizes initially decrease quickly, as a result it is likely that the simple or hyperbolic weighting schemes reflect an intuitive underlying community structure. Analysis of URL metadata (such as TF-IDF weighted importance of web domains within communities) supports this assertion by identifying qualitatively consistent communities formed around geographical or ideological factors.

Our exploration of synthetic networks covers a particular test case. We sample networks with approximately equal mode and community sizes and the same degree distribution for the left and right modes. Future work can expand on our analyses by permitting varying sizes and degree distributions in the modes. We exclude this work here given the combinatorially large search space for the various parameter combinations.

The adjusted Rand index scores in Fig. 4.6 demonstrate that there is merit in using the unipartite projection process to identify community structure present in the bipartite network. Success with this method requires a sufficiently high community

preference p to overcome the influence of the cliques formed by high-degree bipartite nodes on unipartite community detection. If the underlying level of community preference is too low, modularity maximizing community detection produces a poor representation of the bipartite network communities. This problem is exacerbated in degree distributions without a long tail; the existence of high-degree vertices within communities benefits the performance of community detection algorithms, and such vertices imply a long-tailed degree distribution.

It is important to note that compared to a baseline of zero, the modularity results found by community detection on the unipartite projections are deceptive as relatively high modularity is found even in the absence of *any* imposed bipartite community structure. In such circumstances, the projection process is creating local structures which are identified as spurious communities by modularity maximizing algorithms. Despite these concerns, there is evidence that unipartite community detection does recover information about the bipartite network; in the regime of high p, where community preference is strong, dense connections between cliques promote the identification of meaningful partitions through modularity optimization. Considering the standard null model with which network modularity is normally computed, the core assumption of edge independence is violated by the unipartite projection process; projection creates cliques rather than independent edges. It is possible that an adjusted null model which accounts for cliques may facilitate better community detection in projected networks.

Fig. 4.7 shows that the random walk, cosine and Jaccard projections perform similarly to the hyperbolic weighting in many cases but demonstrate some undesirable characteristics. The Jaccard and cosine weightings allow for partitions with very high modularity to be found even when there is no underlying community preference, particularly in the case of the zeta degree distribution. The random walk weighting does not suffer as much from modularity inflation, but requires a greater underlying preference to reproduce the true community structure with the same accuracy as the hyperbolic weighting.

Our experiments allow us to provide a recommendation for which of the seven

projection methods studied is best overall. We frame such a recommendation in the experimental setting where the underlying community preference and structure is unknown and account for accuracy to the underlying structure, modularity of the optimal partition and distribution of community sizes; these factors are all considered with their variance across the ensemble of model realisations. Our results demonstrate that the hyperbolic weighting scheme is the overall best method of the seven studied here, particularly for networks with long-tailed zeta degree distributions (as commonly found in socio-technical systems). As shown in Fig. 4.6a, the adjusted Rand index reveals that the hyperbolic weighting scheme most accurately recovers the bipartite community structure in nearly all cases, after a threshold of sufficiently strong community preference is passed. Hyperbolic weighting has the additional benefits of suppressing the inflated modularity scores common to many of the other methods, and finding meaningful community sizes when a bipartite community structure exists to be found. The hyperbolic scheme also maintains small variance across the ensemble runs, suggesting more robustness to noise in the network. Beyond this optimal method, we also find that the binary and simple weighting schemes give qualitatively and quantitatively similar results in all experimental settings, suggesting that the simple weighting performs no better than an unweighted network.

Overall, this study of how community detection is affected by the edge weighting applied to the unipartite projection of bipartite networks with variable imposed community structure shows that careful thought needs to be given to the application of this approach and the interpretation of results. In terms of accuracy to the bipartite community structure, a useful direction would be to improve algorithms for community detection on bipartite networks directly; we note some recent efforts in this area [210]. However, the projection approach is suitable in many circumstances. If a network arises through a growth process without preferential attachment, modularity maximizing community detection should be used carefully, as in these cases the detected modularity can be very high regardless of the underlying community structure. When the network growth process is driven by preferential attachment

(as is the case in many real-world systems, and social networks in particular), the use of the hyperbolic weighting proposed by Newman [135] generally finds the most accurate results to the true community structure. Modularity found on the unipartite projection cannot be thought of as directly representative of the community structure of the bipartite network as there are several weaknesses in the unipartite null model in this context. Future research into an alternative null model that better reflects the properties of projected networks (e.g. [13]) would be of benefit to the wider scientific community given the widespread application of this approach when studying bipartite networks. An alternative direction for future work could apply the methods outlined here to explore how other measures of network structure are affected by edge weighting during the unipartite projection process.

From theory into practice

What does the previous chapter tell us about collective attention?

In the previous chapter, we set out to answer the question of how much impact the choice of weighting applied to unipartite projections of bipartite networks has on subsequent community detection. Seven candidate weighting schemes were chosen and compared across different degree distributions. While answering this question does not tell us anything about collective attention directly, it improves our understanding of a methodological approach that has been widely used to study large bipartite networks (e.g. [52, 164, 197]).

We saw in Fig. 4.3 that the choice of weighting can lead to different results, even in relatively simple networks. Following this, we tested a series of synthetic networks to investigate how the known community structure responds to projection. We sought to understand the crucial question of how well do the communities on the projected networks reflect the true community structure imposed on the synthetic networks (Figs. 4.6, 4.7). Here we found that a critical threshold for community preference must be passed in order to see any similarity between the bipartite and unipartite communities. We also compared the modularity of the detected partition under each weighting scheme and saw that the projection process leads to an unexpectedly high modularity in networks with low community preference. The final characteristic studied was expected community size as we want to better understand how quickly communities converge to the true size.

Through all of these metrics, we seek to find the weighting that produces the

greatest similarity between the unipartite community structure found in the projected network and the underlying community structure imposed on the bipartite network (where the preference is strong enough to produce true communities); suppresses the inflated modularity when the underlying preference is small; and quickly converges to communities of appropriate size for the imposed communities. By these criteria, we find that the hyperbolic weighting scheme proposed by Newman [135] performs the best, particularly for networks with heavy-tailed zeta degree distributions. Since such heavy-tailed degree distributions are common in social networks, and especially social media, it is this weighting scheme that we will prefer going forward. The hyperbolic weighting scheme also behaves in a useful manner by mitigating the influence of high degree nodes on the weight of any single edge in the projection, thereby allowing typical nodes (by degree) to remain relevant on the local scale compared to more highly connected nodes.

The previous chapter showed an example of a real-world network of users and the URLs they share on Twitter, and demonstrated that the boundaries and relative sizes of communities vary under the chosen weighting schemes. Now that we have determined that the hyperbolic method is preferable for our circumstances, we can explore this dataset of tweets on the topic of climate change in more detail.

How does this inform our upcoming work?

Having developed the knowledge of which projection weighting scheme leads to the most accurate detection of the true bipartite communities, we are now in a position to study a collective attention event using these techniques. In particular, we are interested in understanding how the highly polarised climate change debate responds when collective attention causes many more people to join the conversation and whether the echo chambers seen in previous work (e.g. [196]) are affected.

On 1st June 2017, President Trump announced his intention to withdraw the United States from the Paris Agreement on climate change, an international agreement "to strengthen the global response to the threat of climate change by keeping a global temperature rise this century well below 2 degrees Celsius above pre-industrial

levels and to pursue efforts to limit the temperature increase even further to 1.5 degrees Celsius" [5]. Social media conversations on the terms "climate change" and "global warming" subsequently increased dramatically as users express their opinions and engage with one another after the subject is catapulted to the forefront of the news cycle.

In the next chapter we study collective attention to the climate change debate in the time before, during and after the announcement by President Trump to understand how the community structure of URLs shared in tweets changes over time. We will use text gathered from many of the shared links to explore how the community structure we detect is related to political bias and attitudes to climate change, and how these opinions are linked in the web pages shared. Finally, we address the question of how the increase in media attention to the topic affected the existing network structure, and in particular how increased participation by users is reflected in network attributes.

Chapter 5

Ideological biases in social sharing of online information about climate change

Abstract

Exposure to media content is an important component of opinion formation around climate change. Online social media such as Twitter, the focus of this study, provide an avenue to study public engagement and digital media dissemination related to climate change. Sharing a link to an online article is an indicator of media engagement. Aggregated link-sharing forms a network structure which maps collective media engagement by the user population. Here we construct bipartite networks linking Twitter users to the web pages they shared, using a dataset of approximately 5.3 million English-language tweets by almost 2 million users during an eventful seven-week period centred on the announcement of the US withdrawal from the Paris Agreement on climate change. Community detection indicates that the observed information-sharing network can be partitioned into two weakly connected components, representing subsets of articles shared by a group of users. We characterise these partitions through analysis of web domains and text content from shared articles, finding them to be broadly described as a left-wing/environmentalist group and a right-wing/climate sceptic group. Correlation analysis shows a striking positive as-

sociation between left/right political ideology and environmentalist/sceptic climate ideology respectively. Looking at information-sharing over time, there is considerable turnover in the engaged user population and the articles that are shared, but the web domain sources and polarised network structure are relatively persistent. This study provides evidence that online sharing of news media content related to climate change is both polarised and politicised, with implications for opinion dynamics and public debate around this important societal challenge.

5.1 Introduction

In spite of scientific consensus on the causes and primary effects of climate change, it remains a controversial topic in public and political discourse. Surveys have long shown substantial variation in public beliefs around climate change (for example [112, 128]) and the level of polarisation between individuals supporting and opposing action to mitigate anthropogenic climate change has been growing [54]. Media coverage of climate science and the frames used to present the information can have an important impact on public perceptions and willingness to take action [63], present different motivations and calls for action [151] and influence the accuracy and longevity of reproduced messages [44]. Recent work has shown that media effects vary depending on existing political biases [205]. Understanding the media landscape around climate change is of key importance in mapping public engagement with the issue and support for political actions to confront it.

Assessing which people are exposed to what information is fundamental to any study of the effects of media on public understanding and opinion. The disruptive impact of online media has transformed the media environment, radically altering the diversity of content people encounter as well as the exposure process itself. Individuals are faced with a wide range of media options (both social and traditional), new patterns of exposure (selected by the end user or driven by their social network) and increased production of user-generated content [179]. Previous work in this area has focused on the effects of incidental exposure on media awareness (e.g. [62]) and the diversity presented by online recommender systems (e.g. [131]). Such efforts are

hampered by the diversity of online platforms, the rapid pace of their development and the obfuscation of the algorithmic processes they follow, and as a result no universal understanding of exposure effects is possible. Whether an individual is consuming the news online from a legacy media organisation, or producing and consuming information on social media, the fundamental dynamic of communication exposure and influence is that of network formation [36], based on creation of new relationships between users and media content by a variety of means (such as web-browsing and social information-sharing). Online media exposure creates a network that links sources and consumers of content (nodes) via their interactions (edges), requiring a network perspective for its proper understanding.

This study aims to describe patterns of sharing online media content about climate change. While a complete record of users' exposure to digital content is only possible via accurate tracking of web browsing histories, media engagement can be inferred from the content users share on social media. Sharing a web article requires action by the user, and causes it to appear on the social media feeds of friends and followers, as well as contributing to aggregate trends, often advertised by social media platforms. This is used to indicate a significantly higher level of engagement than exposure. Social sharing of content instantiates a promotion mechanism, increasing the visibility of any content across a user's social network and likely indicating that the sharer agrees with, or approves of, the content. These factors, along with the increasing use of social media as a source of news [167], mean that study of digital media sharing can provide insights into how information is propagated online, including important contemporary issues like climate change. In particular, Weaver et al. [187] shows that network analysis of such propagation patterns can reveal meaningful social structures of news engagement and consumption around political events.

Here we operationalise our study of online information-sharing around climate change by examining link-sharing on Twitter. User posts (tweets) referencing climate change and containing links to web content (URLs, which are often rendered into news blurbs or images in the Twitter client) were collected from the Twitter

social media platform via its public API. This dataset was used to construct bipartite networks linking users to the digital media they shared. Analysis of network topology finds strong community structure, supplemented by comparison of source domains and textual content of articles shared within each community. Exploration of detected communities identifies strong ideological polarisation within the news-sharing network where users are segregated by divergent opinions, rather than an ongoing process, an alternative definition which is important in other contexts [53]. Overall the results indicate highly polarised and politicised engagement with web content around climate change, with largely segregated and ideologically biased communities receiving information from different media sources. We demonstrate that the observed correlation between political views and climate change beliefs (e.g. [54]) extends to online information sources and their shared readerships. For the first time, we track network structure over a 7-week period, including a disruptive media event (the announcement of the withdrawal of the USA from the Paris Agreement on climate change), to show the persistence of the observed polarisation and politicisation of media sharing related to climate change over time and with varying background levels of public interest.

Social media data has been used to study several aspects of public opinion around climate change, including attitudes towards and engagement with climate change mitigation strategies [66], media framing of the leaked "Climategate" emails on YouTube [151], and the spreading of calls for collective action at the COP15 conference [166]. Networks are an intuitive representation, and come with a host of analytical tools to understand the shape of online discussions around climate change; for example Elgesem et al. [56] explore the network of hyperlinks between blogs, while Williams et al. [196] study the structure of follower, retweet and mention networks on Twitter. In both of these studies, user communities manifest as densely interconnected clusters with similar characteristics. These communities are highly polarised, such that each community is well described by a single viewpoint, with few moderate voices. Similar patterns have also been observed for online political discourse (e.g. [6, 15, 41, 45, 46, 190]). This pattern of opinion polarisation and seg-

regation has important implications for opinion change and the likelihood of global consensus [174]. Polarisation in online social media is most frequently studied in the political sphere, especially for two-party political systems with an ideological split along a left-right axis. However, the phenomenon also extends to the competing opinions around climate change, which are often simplified as a debate between environmentalists (supporting the scientific consensus and promoting action) and sceptics (doubting or opposing the consensus and need for action), notwithstanding the diversity of viewpoints and representations of this complex issue [145]. These previous network-based studies generally treat datasets as single snapshots, along with the implicit assumption that the phenomena under study varies slowly. The intervals studied range from months to years, but by choosing such a timescale it is possible to overlook the changes that social networks can experience in shorter periods.

The pattern of polarisation in both political and climate change contexts is often associated with the existence of echo chambers in the social media ecosystem, whereby users choose to associate with people and news-media sources which conform to and reinforce their existing beliefs [196]. Echo chambers have been proposed to contribute to the spreading of misinformation [50], political networks of environmental actors [94], exposure to political information on social media [190], and online content around climate change (e.g. [56,82,196]). In this work we focus on the structural phenomenon of echo chambers but other scholars have looked at how they are linked to psychological processes such as confirmation bias (e.g. [51]).

Previous studies of information-sharing around climate change have focused on the prominence of different sources. Newman [142] analysed the tweets and information sources shared alongside the release of the IPCC AR5 WG1 report, finding a focus on the public engagement with science, and a dominance of mainstream media sources. Segerberg and Bennett [166] examined the breakdown of different link sources used alongside calls to collective action at the COP15 conference. Kirilenko and Stepchenkova [99] studied the URLs shared on Twitter over the course of one year in five different languages, finding that by country, the US dominated total

tweet counts, and a mix of traditional media, activist and sceptic sites were shared.

Polarisation is a common observation in the climate change debate. Notable studies such as Dunlap et al. [54] have shown that the polarisation effect in climate change opinion has grown between 1997 and 2016. An impact of polarisation can be observed in the frames used to discuss climate change, such as by Jang and Hart [93], who analysed the themes present in Twitter's climate change debate across two years. They found differences in the terminology used by opposing groups, with Republican-leaning states in the US using *global warming* in preference to *climate change*, and often using hoax frames to cast doubt on the scientific consensus. O'Neill et al. [144] found that this trend extended to the media coverage of the publication of the IPCC AR5 working group reports. By studying the frames used in newspaper and television broadcasts, they found clear preferences for different frames amongst the various news organisations. Some work has looked at countering the growing levels of polarisation, including Zhang et al. [205] who study the effect of clarifying messages on accuracy over the perceived levels of consensus among climate scientists. Among their experimental group, exposure to the clarifying message lead to more uniform accuracy around the scientific consensus through greater impact in the areas of lower baseline belief.

In polarised systems, different points on the opinion spectrum often display different behaviours. Schuldt et al. [165] compared the usage of the terms *climate change* and *global warming* across the websites of a series of think tanks. Right-wing think tanks were more likely to use *global warming*, with the opposite trend observed in left-wing think tanks. These findings of content differences from polarised sources extend beyond the climate change debate. Further analysis of climate sceptic organisations and their funding sources by Farrell [61] found shared sources of funding across many of them. Beyond the topic of climate change, Freelon et al. [70] present an overview of how the different online activism strategies of left- and right-wing groups manifest different types of content and audiences. They highlight a key difference in the perceived strategies of the different groups. Left-wing groups target "hashtag activism" leading to social promotion of movements whereas right-wing

groups engage much more readily with sympathetic media organisations to promote their messages and goals. Freelon et al. also recognise a similar trend in group coherence where right-wing groups are tighter when compared to left-wing groups formed from a loose coalition of multiple issue-led groups. Considering this coordinated funding of multiple groups and engagement with media organisations, it is to be expected that climate sceptic messaging is likely to be more consistent than competing environmentalist information.

Along with climate change, studies of information-sharing on social media have embraced a diversity of different topics; Starbird [172] studied the spread of misinformation around mass shooting events in the United States, finding a cluster of alternative news sites separate from mainstream media sources. Arif et al. [11] and Del Vicario et al. [50] study the spreading behaviour of rumours and misinformation on Twitter and Facebook respectively and classify observed trends. Schmidt et al. [164] analyse a network of news-related pages on Facebook, where two pages are connected if they have posts that are liked or commented on by the same user. Their cluster analysis reveals a highly polarised structure, as seen in a number of other contexts (e.g. [46]). Williams et al. [197] also study the community structure of political news-sharing via Twitter, finding communities characterised by both geographical and political factors. Weaver et al. [187] examined information-sharing on Twitter during the UK General Election in 2015, showing strong community structure explained by ideological, geographical and topical preferences. Each of these studies should be considered in the context of the typical sharer and the information they are exposed to. Not all users on social media are exposed to the same information, and typically they are exposed to information in which they have already shown an interest through their own decisions of which users to follow, along with algorithmic filtering effects. Sharing information requires action on the user's part and as such those who choose to share are likely part of a more highly invested subset of the users exposed to a certain piece of information.

The rest of the paper proceeds as follows. Section 5.2 details the method of data collection and preparation, along with the techniques used to construct and

analyse the information-sharing networks. Our main results follow in Section 5.3, including network analysis and characterisation of community structures, and finally Section 5.4 provides a thorough discussion of the main findings of this study and places them in the wider context of media effects on the climate change debate. Additional data and visualisations can be found in the Supplementary Information.

5.2 Methods

5.2.1 Tweet collection and pre-processing

This analysis uses seven weeks of Twitter data collected from the Streaming API [4]. Tweets containing the strings *climate change* or *global warming* were collected between 2017-05-10 and 2017-06-27, giving an initial dataset of $5,320,400$ tweets by $1,975,593$ users. Inspection suggests that most of the content comes from the US and the UK based on the presence of many news sites and other domains with a primary focus on these two countries.

The collection period captures a key event in the unfolding climate change narrative, the announcement on 1st June 2017 by then-US President Donald Trump that the USA would withdraw from the Paris Agreement on climate change mitigation. This event caused a large spike of activity on Twitter around the topic of climate change illustrated by Fig. A.1. To study this event in the context of a longer period of "normal" activity, the collection spans seven weeks centred on the week of greatest activity. We separate the dataset into seven one-week intervals to account for the potential weekly periodicity in social media usage and give sufficient sampling density for robust network analyses.

This study focuses on the digital media engagement and sharing behaviour of Twitter users discussing climate change. To capture this behaviour, the dataset is filtered to remove all retweets (where a user reposts an original tweet) and quotes (where a user reposts an original tweet with their own commentary prepended). The purpose of this filter is to focus on original tweets, which are assumed to be the strongest available measure of user engagement with content; the low effort cost of

Week	Original Tweets	Unique users
1	52,737	27,499
2	60,922	31,769
3	78,353	38,683
4	209,637	96,506
5	67,234	36,620
6	65,248	34,625
7	58,699	28,536

Table 5.1: Weekly tweets and unique users in our filtered dataset. Week 4, in which the Paris Agreement announcement was made, includes many more tweets than any other week, accounting for 35% of all the tweets studied.

retweeting and quoting means that user engagement cannot be inferred as clearly from these tweet types. The remaining tweets are further filtered to retain only those which contain an embedded link (URL) to web content. These links are the digital media items shared by the tweet author, including news articles, blog posts, videos and other content. Such tweets can be composed by manual insertion of a link or clicking the 'share' button often presented alongside online news content. Application of all the filters leaves a dataset of 592,830 tweets containing URLs by 195,134 users, broken down over the 7-week study period in Table 5.1. Most notably, the centre point of the dataset in Week 4 includes the US intention to withdraw from the Paris Agreement, producing a dramatic surge in Twitter activity.

5.2.2 URL validation

Every URL found in the tweet dataset was validated in December 2017 to ensure that it was a working link to an identifiable item of online content. This validation step was necessary for a number of reasons. Firstly, each URL must be accessible to allow subsequent analysis of content. Secondly, typographical errors by tweet authors are sometimes incorrectly interpreted by the Twitter API as URLs, such as periods followed by alphanumeric characters (*hello.world*). Thirdly, the validation process handles shortened URLs, which represent the majority of URLs in tweets. Many Twitter users make use of third-party URL shorteners, services which create a redirect to a given long URL from a shorter version that can then be used to reduce the characters needed to embed a URL in a tweet (this is unnecessary since Twitter

includes its own shortening service that reduces any URL to 23 characters). URL shorteners are also sometimes used to conceal the target URL from anyone who may potentially click on it. These short links must be traced to their destination both to ensure they point ultimately to a valid resource, and also to be treated alongside other links to the same destination.

Each URL is resolved individually, marking as valid only those that return HTTP status code 2XX immediately or through a small number of permanent redirects ($301 \rightarrow 301 \rightarrow ... \rightarrow$ 2XX status codes). The 2XX status codes indicate various successful outcomes to an HTTP request.

Many valid URLs included modifiers inserted for tracking and metadata purposes, without altering the destination page. This can lead to many URLs pointing to the same resource, so to minimise the possibility of considering these as distinct URLs, all schemes and queries are removed from the resolved URLs. A small number of domains were adversely affected by this. YouTube and other Google services were the most prominent, largely due to the structure of their links. As a result, our process does not distinguish between different YouTube videos within the networks.

The validation process results in a mapping between each raw URL (as given in a tweet) and the validated final destination after any acceptable redirect paths. Any tweet containing at least one unvalidated URL is removed from the dataset. Finally, all users who tweeted more than 50 times within a given week are removed; this threshold is found to be a sensible limit to mitigate the impact of automated accounts, especially news aggregator accounts. Across the seven weeks 271 unique accounts were affected, leading to the removal of $66,892$ tweets (approximately 1% of the total dataset). This left $245,446$ tweets by $113,154$ users sharing $54,462$ distinct, validated, URLs across all of the seven one-week windows.

5.2.3 Network construction

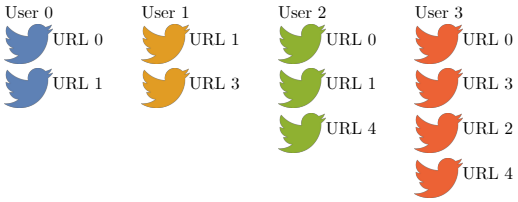
Bipartite networks, containing two node classes representing users and URLs, are constructed by creating an edge $i \rightarrow j$ whenever user i shares URL j, illustrated by Fig. 5.1. Multiple shares of URL j by user i increment the edge weight so that the

final bipartite edge weight $w(i, j)$ represents the number of times user i shared URL j.

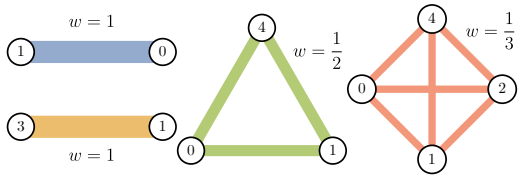
Next a unipartite projection is produced from the bipartite network, whereby a network of only URL nodes is created, where edges encode the number of users that shared the connected pair of URLs. The network construction process is illustrated in Fig. 5.1. The URLs shared by a single user (Fig. 5.1a) form a fully connected clique of nodes representing the sharing pattern for that user (Fig. 5.1b). The whole unipartite network is then constructed by composing the cliques for all users (Fig. 5.1c). The projection allows the use of efficient unipartite network analysis algorithms and focuses the analysis on the relationships between URLs. The unipartite network of URLs represents the collective pattern of sharing online content of the targeted Twitter user population. Statistics for the seven one-week networks created are given in Table A.1.

We restrict our analyses to the giant component to avoid the issue of network fragmentation and limit its impact on community detection. In all cases the giant component consists of around two thirds of all URL nodes in the network. One important design decision made in this process is how user shares are encoded into unipartite edge weight; different weighting schemes significantly impact the performance of community detection algorithms. User contributions to unipartite edge weights are scaled by the factor $1/(k_i - 1)$, where k_i is the total number of article shares by user i (including repeated links), such that the total unipartite edge weight contribution by each user is $k_i/2$ (as proposed in Newman [135]). This hyperbolic weighting scheme was found to allow robust recovery of community structures after projection [34]. Without this weighting factor, a user's edge weight contribution to the projected network is quadratic in the number of URLs shared, since each user creates $k_i(k_i - 1)/2$ edges in the projection, causing users who share many articles to quickly dominate the network and subsequent analysis.

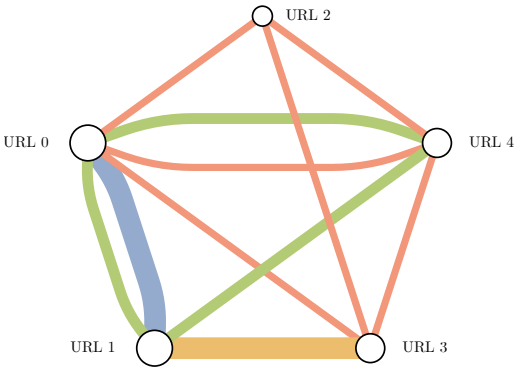
Computing the projection requires knowledge of the biadjacency matrix B of the giant component, where the rows represent users and the columns represent URLs; and a diagonal matrix D such that $D_{ii} = 1/(k_i - 1)$. Given these matrices, the



(a) Over each week in the study period, users share a number of URLs.



(b) Cliques of connected URLs are produced for each user.



(c) The projected unipartite network is a summation of users' cliques.

Figure 5.1: **Schematic diagram of the bipartite network construction and unipartite projection.** Each user is connected to the URLs they share in the study week. The unipartite projection creates edges between two URLs whenever they are shared by the same person. Multiple edges in the projection indicate that multiple users have shared the pair of URLs, and this information is tracked by edge weights. A user's edge contribution to the projection increases quadratically with the number of URLs they share, potentially leading to a dominance of highly active users in the unipartite projection, for example User 3 (red) in the projection; this is handled by a hyperbolic weighting scheme (see Section 5.2.3).

adjacency matrix for the weighted user projection $P = B^T DB$ and hence the weight of an edge between URLs i and j in the projection

$$w(i,j) = \sum_{u \in \text{Users}} \frac{w(u,i)w(u,j)}{deg(u) - 1}. \tag{5.1}$$

5.2.4 Community detection

Community detection was used to find clusters of densely interconnected nodes in the network, which in this context represent sets of URLs which were seen by a group of similar Twitter users. Community detection [140] is a means of algorithmically identifying such clusters in a given unipartite network. This study used a greedy algorithm proposed by Clauset et al. [42] that partitions nodes into communities

such that the modularity of the partition is optimised; modularity measures the proportion of edges within communities relative to the proportion of edges between communities. If a community assignment is significantly better than random, a modularity score between 0.3 and 0.7 is typically observed [140]. Modularity scores for each network are given in Table A.1.

5.2.5 Content analysis

Page content was collected for each validated URL using Diffbot [1], an online service for extracting the constituent parts of an HTML document and presenting them in an easily analysed format. In some cases, Diffbot only identifies the title and cannot automatically detect the content of a webpage. To mitigate the impact of this, the title is used as the page content for such URLs. This substitution affected less than 5% of the URLs in each week, except Week 4 which required substitution in 5.9% of the URLs. Additionally, a small number of domains were incompatible with the Diffbot API. This accounts for around 4% of all URLs in the giant components of any week and mostly arises due to page formatting or timeout issues.

Online content was analysed quantitatively by calculating term frequency-inverse document frequency (TF-IDF) scores, treating each URL as a document and the set of all URLs in the whole network as the corpus. TF-IDF analyses aim to identify distinctive or important tokens (usually words) in a document, based on their frequency in the document relative to their frequency across the corpus. Three kinds of token were studied in this way, in three separate analyses: web domains, page content unigrams, and page content bigrams. Web domains were extracted for each URL to permit analysis of the sources of content shared on Twitter. Words and bigrams (two-word phrases) were extracted from the web pages associated with each URL to allow large-scale analysis of the topics within the content. Before calculating the TF-IDF vectors for page content, the Snowball stemmer [152] was applied to each content token. Each stemmed token is mapped back to the most common token that maps to it for ease of reading, e.g. if fisher, fished and fishing all appear once and fishes appears twice then the final representation of these tokens is fishes.

TF-IDF score vectors for each URL are calculated and represent the frequency of tokens in a document relative to their frequency across the corpus. For token t in document d we have

$$\text{TF-IDF}(t, d) = tf(t, d) \left(\log \left(\frac{1 + n}{1 + df(t)} \right) + 1 \right), \tag{5.2}$$

where n is the number of documents in the corpus, $tf(t, d)$ counts the frequency of t in d and $df(t)$ counts the number of documents in the corpus which contain t. To characterise each community found in the sharing network, the TF-IDF vectors for each URL in a community were summed to obtain aggregate scores. In these community-level comparisons, a high TF-IDF score for a community means that the given token appears more frequently in the URLs in this community, when compared to other communities. In each case, tokens which occur in fewer than 50 URLs, or more than 50% of all URLs in a given week are rejected, such that very common or very rare tokens are omitted. We also removed a set of common stopwords including *trump*, *paris* and *agreement*. In principle, n-grams can be studied with any value of n, but beyond $n = 2$ token frequency is generally too low to be useful.

5.2.6 Ideological coding of source domains

To examine ideological positioning of popular source domains along axes of political leaning (from left/progressive to right/conservative) and climate scepticism (from environmentalist to sceptic), 62 of the 75 most commonly shared source domains across the entire study period were manually coded for ideological bias (listed in Table A.4). Ideology expressed in articles from each domain was graded on a three-point scale for political opinion (Left-Neutral-Right) and climate change opinion (Environmentalist-Neutral-Sceptic), with an additional null rating (Unclear) added to both scales for cases where no clear ideology was seen.

A team of six human coders used this scale to independently score text extracts of articles/content from the most commonly shared domains. For the 75 domains with the most shares across the seven-week period, we sampled the page content for five most shared articles (using Diffbot to extract clean text, as described above).

If fewer than five articles from a popular domain were available in the dataset, then all were included. In some cases, the page formats were incompatible with the Diffbot API, returning no content for 10 domains (see Table A.4 for the excluded domains). We additionally excluded the social media sites *Twitter*, *Wordpress* and *Reddit*; these do not have editorial control and therefore lack a unified ideological position. This left a set of 62 domains to be coded. Each extract consisted of up to three complete paragraphs (of at least 30 words) from the linked web page text. To limit any subjectivity arising from the coders' personal perspectives, each extract was anonymised (i.e. source domain and author information were removed) when presented for coding.

Coders were provided with the following definitions to help make their assessments:

- *Left:* A left-wing stance can be characterised by the promotion of state benefits and services, public investment in and regulation of private businesses, increased taxation of corporations and high earners, and support for workers and trade unions.

- *Right:* A right-wing stance can be characterised by promoting low taxation and minimising the interference of government in personal and business lives. Public investment is minimised, in favour of allowing market forces to control growth and provision of services.

- *Environmentalist:* An environmentalist stance supports the scientific consensus on anthropogenic climate change and promotes immediate action by governments and individuals to mitigate the future impacts.

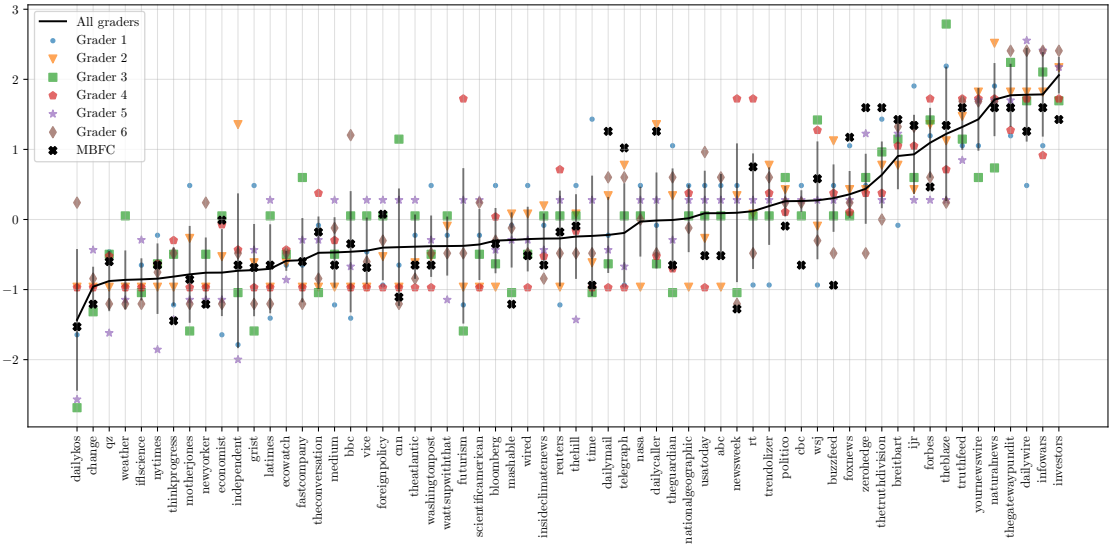
- *Sceptic:* A climate sceptic stance opposes the scientific consensus on anthropogenic climate change. Such opposition varies from questioning the existence or causes of climate change, to opposing efforts to mitigate its impacts.

Each coder assigned a score to each article extract independently. That is, six coders generated up to five scores for each of 62 domains (in practice, this amounted to 1,698 scores out of a maximum of 1,860, with each domain receiving up to 30 scores). Each

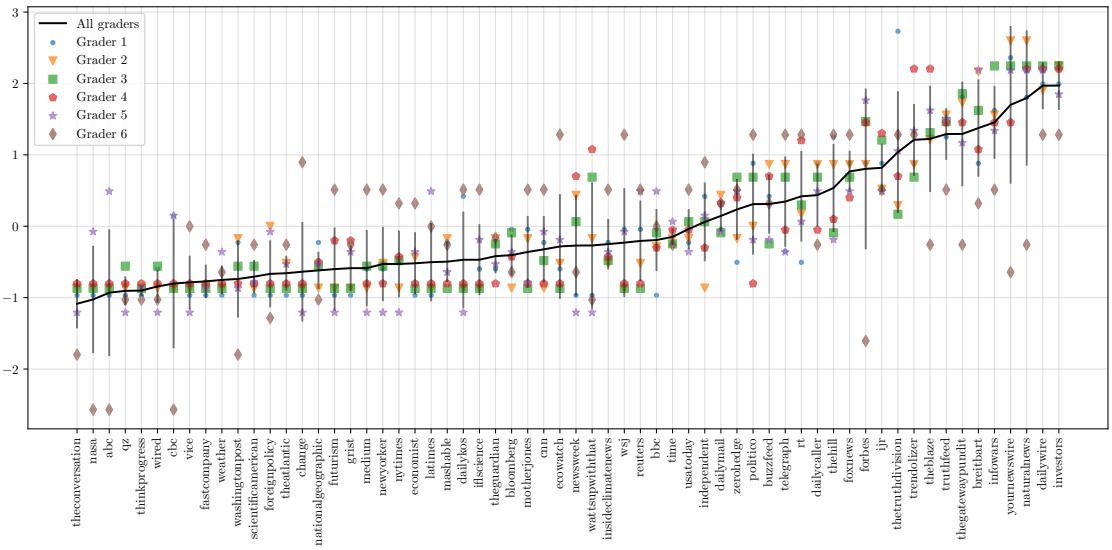
article score is an integer denoting the level of ideological bias the coder observes in that article (-1/0/1 for left/neutral/right or environmentalist/neutral/sceptic, in order, with unclear scores ignored; see Table A.3). The article scores assigned by each coder were then averaged for each domain to determine an overall domain score for that coder. Each domain score is a real-valued number in the range $[-1, 1]$; there are 6 scores for each domain, one from each coder.

Since the assignment of ideological bias is somewhat subjective, we performed an adjustment to the domain scores assigned by each coder. This adjustment normalised each domain score by subtracting the mean domain score for that coder across all 62 domains, then dividing by the standard deviation. This process represents domain scores as z-scores and normalised for subjective bias of individual coders. Finally, we average the normalised domain scores across all coders to find a single domain score. This process was applied for both political and climate change ideology. Since judgment of climate change ideology relies on some knowledge of climate change (e.g. the scientific consensus) the coders were recruited from the postgraduate research community in sciences at the University of Exeter.

Full results for the coding of each domain are presented in Fig. A.2. To test the reliability of our coding exercise we compare our grades to the domain assessments by Media Bias/Fact Check [2] (MBFC), a fact-checking and media bias site. We translate the MBFC ratings to our own scale and apply the same z-score normalisation, then calculated the correlation of the political bias scores by our coders against those from MBFC for those domains which have been rated by MBFC. We find strong positive correlation (Pearson's $r = 0.8$, $p < 10^{-6}$, $n = 49$) demonstrating that the political coding process is consistent with the expert assessments provide by MBFC. No equivalent external assessments are available for climate change ideology, but the robust performance of our coding process for political bias gives confidence in the methodology. See Figs. 5.2a and 5.2b for the average results across domains and coders for political and climate change bias respectively.



(a) Political ideology determined by the coders.



(b) Climate change ideology determined by the coders.

Figure 5.2: **Average domain ideological positions assigned by each coder and the overall average across all coders.** Values should be thought of as standard deviations from the mean. Vertical bars indicate ± one standard deviation across the coders.

5.2.7 Comparisons over time

To measure the change in patterns of sharing climate media over time, the pairwise similarity of the sets of users, articles (URLs) and source domains was calculated for the seven weeks in the study period. An asymmetric similarity measure is defined to give an indication of persistence between weeks and identify influxes of new participants or content. This expression compares weeks A and B by the fraction of the users, u, links, l, or domains, d, that appeared in week A that also appeared in week B. In order to account for the repeated usage common in online social networks, a measure of how many times a user, URL or domain appears in each of the two weeks is included by using multisets for u, l and d.

$$S^u_{A,B} = \frac{|u_A \cap u_B|}{|u_A|}, \qquad S^l_{A,B} = \frac{|l_A \cap l_B|}{|l_A|}, \qquad S^d_{A,B} = \frac{|d_A \cap d_B|}{|d_A|}. \tag{5.3}$$

Values of the similarity measure fall within the range $[0, 1]$. Values approaching 0 signify that very few of the users (respectively URLs, domains) in week A are also present in week B, whereas values approaching 1 signify that nearly all of the users (respectively URLs, domains) in week A are also present in week B. Note that by design $S_{A,B} \neq S_{B,A}$ for $A \neq B$ in general.

5.3 Results

This section divides our results into three main findings. In the first part we focus on Week 4 of the study period, in which the US withdrawal from the Paris Agreement was announced, to demonstrate the broad ideological polarisation observed in the information-sharing networks. Secondly, we characterise the sub-communities that make up the network, showing that several linked left-wing/environmentalist communities co-exist with a single right-wing/sceptic community. Finally, we look at all seven weeks in the study period to explore how network structure and polarisation change over time.

5.3.1 Climate media sharing is polarised and politicised

We begin by characterising the information-sharing network during the central week of the study period, Week 4, in which the US withdrawal from the Paris Agreement was announced. During this week, 7,496 URLs were shared by 42,113 Twitter users. After projection this produces a unipartite network of 7,496 URLs connected by 107,304 edges indicating which pairs of URLs were co-shared. Fig. 5.3 shows the URL co-share network of the five largest communities by total share count. The layout is determined by the ForceAtlas2 algorithm [92], which groups densely connected nodes together; visual inspection shows a clear partition into two large clusters. Algorithmic community detection reveals further partitioning within the larger cluster of the network, illustrated by the different node colours. By considering the contrasting structures found by the community detection and network layout algorithms, there is evidence of multiple layers of structure within the network. Note that the large yellow node separated from the two main clusters is *youtube.com/watch*. This node is difficult to interpret as it aggregates many YouTube video links, seemingly from all sides of the debate. We cannot verify this with our text-based content analysis.

There is a strong correlation between opinions about climate change and political ideology expressed in shared content. Fig. 5.4 plots the position of the 62 web domains on axes of climate bias and political bias, based on content of articles coded by the panel. Left-right political ideology and environmentalist-sceptic climate opinions are very strongly correlated (Pearson's $r = 0.86$) and very few domains appear in the right-wing/environmentalist and left-wing/sceptic quadrants.

Mapping ideologies/opinions associated with web domains onto the information-sharing network structure shows an association between network position and viewpoint. Fig. 5.5 shows the network diagram from Fig. 5.3 with URL-node colours altered to show biases in the political/climate opinions expressed by their web domains. Fig. 5.5a colours the nodes to highlight the left-right political bias. The left-hand cluster contains predominantly left-wing sites. The right-hand cluster has a high concentration of right-wing sites, but also has many uncoded sites. A sim-

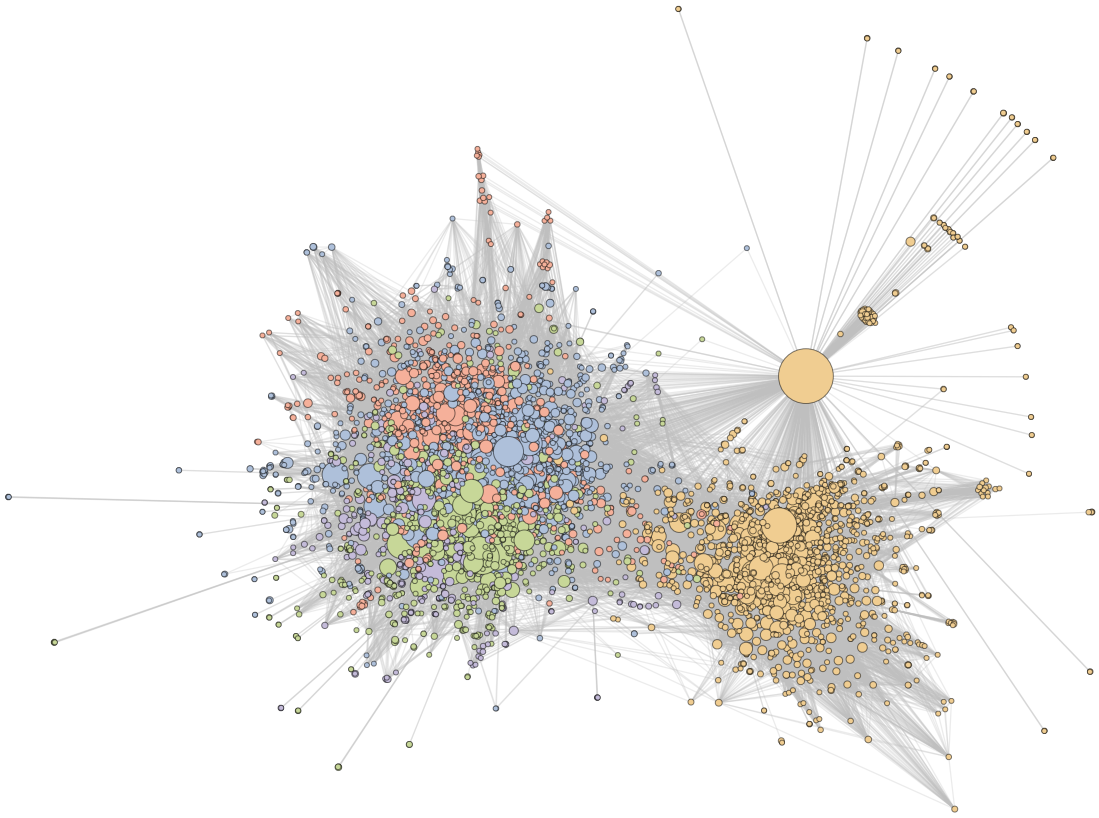


Figure 5.3: **Information-sharing networks are polarised.** Plot shows the URL co-sharing network for the week in which the US withdrawal from the Paris Agreement was announced (Week 4 of the study period). The five largest communities by total share count are displayed (67.69% of 7,496 nodes). Communities $1-5$ are coloured blue, yellow, green, red and purple respectively and node size is proportional to the square root of total share count. Node placement uses a force-directed algorithm [92] which groups densely connected nodes together; this layout highlights two large clusters, with four communities on the left and a single community on the right.

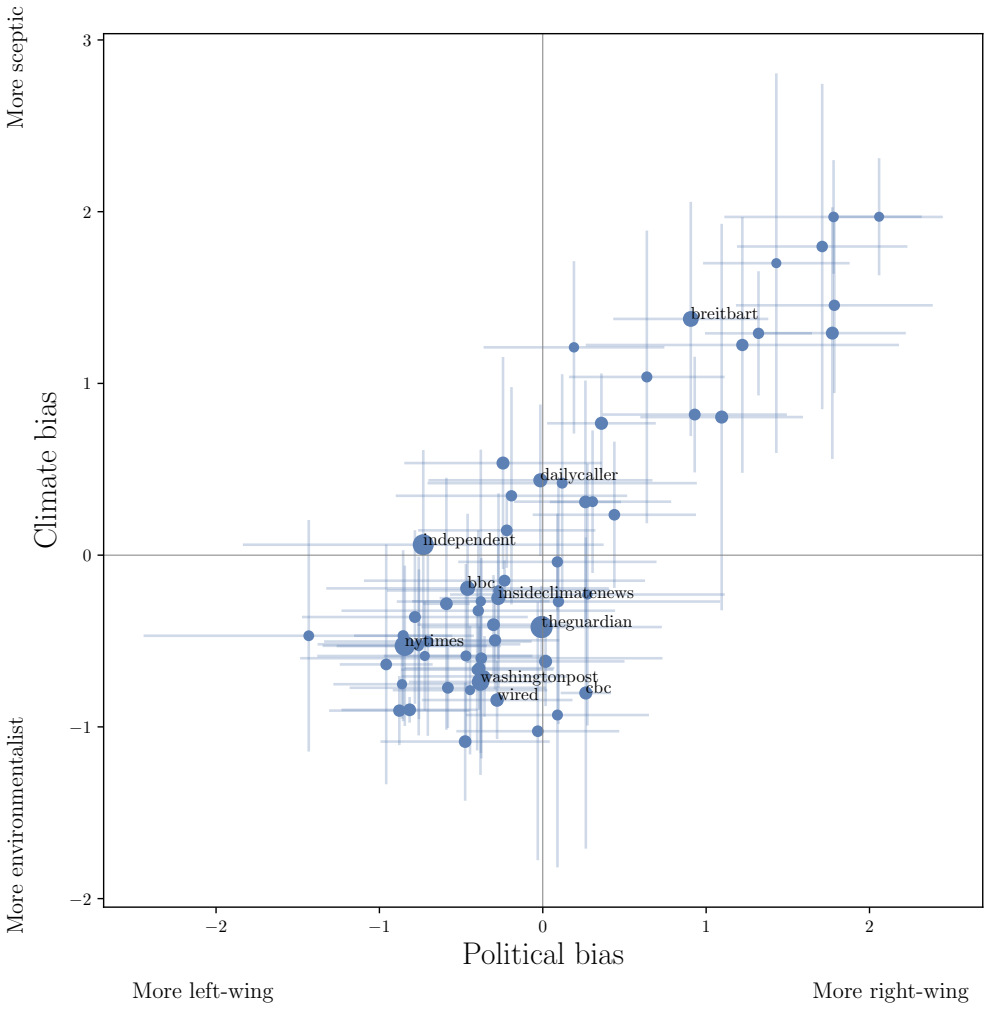


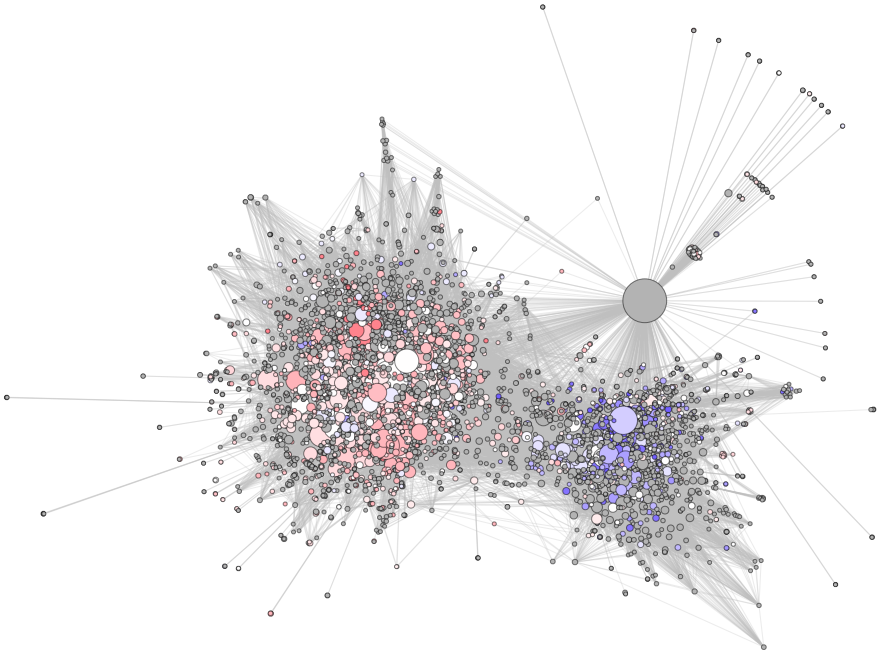
Figure 5.4: **Climate media content is politicised.** Mean political ideology (left-to-right) and climate opinion (environmentalist-to-sceptic) expressed in content from the 62 coded web domains over the six coders (see Section 5.2.6). Point size is proportional to the square root of total share count and lines indicate ± one standard deviation. Labels are shown for 10 most frequently shared domains in the coded list.

ilar pattern can be seen in Fig. 5.5b, which colours nodes by climate change bias, with the left-hand cluster predominantly environmentalist and the right-hand cluster containing most of the sceptic domains. Taken together, these findings demonstrate significant strength of polarisation and politicisation in information-sharing about climate change on Twitter, with two large clusters of users and information sources, broadly characterised as a left-wing/environmentalist group and a right-wing/sceptic group.

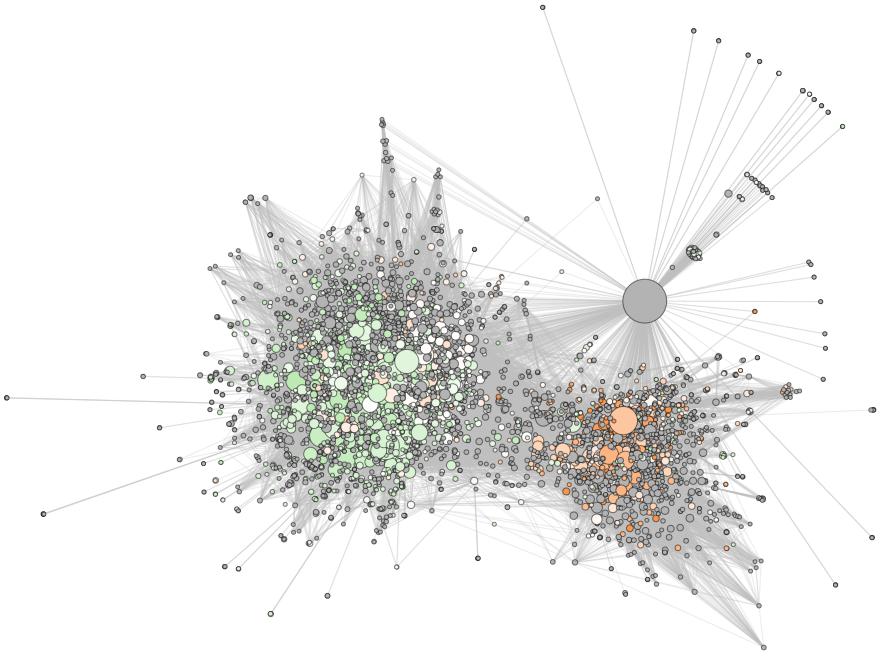
5.3.2 Characterisation of information-sharing communities

The main source domains and indicative content of media articles shared within the five largest communities in Week 4 can be seen in Fig. 5.6, in which the radii of the circles show the relative sizes of these major communities. Here Communities 1, 3, 4 and 5 are communities within the left-wing/environmentalist cluster in the information-sharing network, whereas Community 2 is the single community in the right-wing/sceptic cluster. Fig. 5.6 uses a TF-IDF weighting scheme for each token, such that the prominent tokens are those that are characteristic of a particular community when compared to the network as a whole. Source domains are shown in Fig. 5.6a and content is shown in Figs. 5.6b (unigrams) and 5.6c (bigrams), to allow a characterisation of the broad themes in each community in terms of geographic focus, political or climate science biases and key subjects of interest. Table 5.2 summarises these communities using the data from Fig. 5.6.

Looking at source domains (Fig. 5.6a), different geographic and political biases can be inferred based on earlier analysis of domain ideology. Community 2 features predominantly right-wing sources, whereas Communities 1 & 3-5 contain content from left-wing sources. Terms such as *decision* and *withdraw* in each of the communities show that the Paris Agreement announcement is a major topic of conversation during this seven-day period, even spanning the ideological/geographic divisions illustrated by Fig. 5.6a. The left-wing communities are reasonably similar to each other, while the right-wing community is unique in its mention of previous US political figures such as *Obama* and *Gore* and use of scientific terminology. The

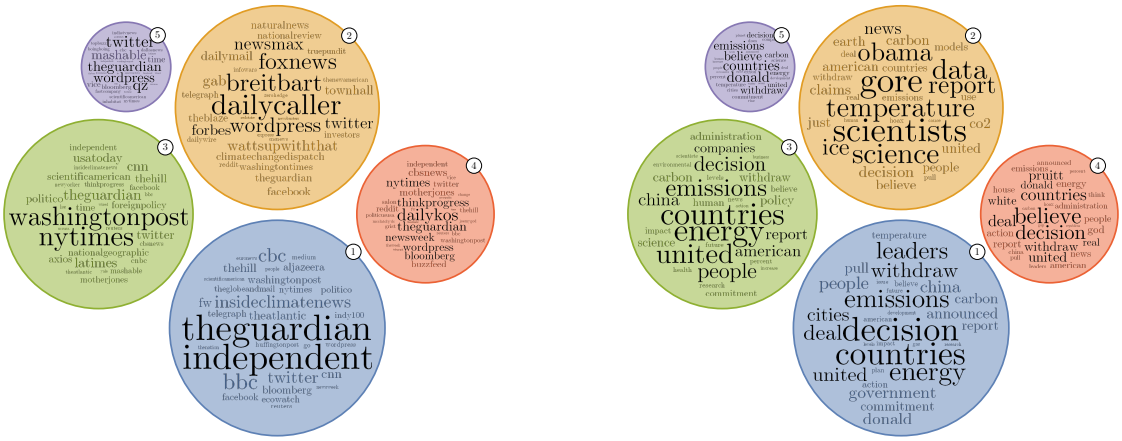


(a) Political bias.



(b) Climate change bias.

Figure 5.5: **Network clusters are ideologically biased.** The two large clusters within the URL co-sharing network for Week 4 shown with URLs coloured by: (a) the average political bias of their source domain; and (b) the average climate change bias of their source domain. Red denotes left-wing domains, blue denotes right-wing domains, green denotes environmentalist domains, orange denotes sceptic domains. White denotes any domain coded as neutral and domains not coded are in grey.



(a) Source domains for each community.

(b) Unigrams from shared articles.

(c) Bigrams from shared articles.

Figure 5.6: **Source domains and content of media articles shared within the five largest communities in the information-sharing network.** Tokens in each plot are weighted (using TF-IDF) to make distinctive tokens prominent. Circle size is scaled to indicate the total number of URL shares within the community. Terms coloured black are the highest weighted terms required to reach 15% of the total weight in a community. For visual clarity, each stemmed token is represented by the most common token that maps to it (or pair of tokens for bigrams).

Community 1	This is the largest community and is dominated by mainstream media outlets, such as *The Guardian* and *The Independent*, mostly from the UK. This community includes mainstream news reporting and discusses the Paris Agreement on climate change almost exclusively, focusing on the consequences of President Trump's decision and any international responses.
Community 2	This community includes many right-wing sources, including alternative news sites such as *Breitbart* and *The Daily Caller*. Social media sites are also prominent in this community as *Facebook*, *Twitter*, and *Gab* all appear, suggesting that this group captures attempts to use Twitter to re-share content from other social platforms. Some established media outlets such as *The Daily Mail* and *Fox News* are present, but are less focused in this community than alternative news sites. This community also discusses the Paris Agreement decision made by President Trump, as well as certain aspects of climate science. This is the only community to focus on former US politicians *Obama* and *Gore* and the phrase *global warming* dominates the bigram cloud (Fig. 5.6c).
Community 3	This community consists of many mainstream domains revealing a US focus. The most prominent domains here are established mainstream media sources such as the *Washington Post* and *New York Times*. As with Community 1, the Paris Agreement on climate change is a key topic of interest.
Community 4	This community includes a number of alternative and smaller news media domains with a mostly left-wing bias, such as *Daily Kos* and *Mother Jones*, amongst established mainstream news sources. The content here is similar to that of Community 1 and Community 3, but also references then-US Environmental Protection Agency Administrator Scott Pruitt and Michigan congressman Tim Walberg for their comments around climate change.
Community 5	This community is comprised of a mix of social media, news and commentary sites. Again, the Paris Agreement decision is a focus, with additional framing around global consequences and opinion-pieces on the decision.

Table 5.2: Characterisation of the five largest communities in Week 4.

bigrams mostly confirm the findings in Fig. 5.6b, but also highlight differences in terminology, e.g. greater prominence of *global warming* in right-wing Community 2. Fig. 5.6 and Table 5.2 demonstrate that there is variation in the geographic scope of the communities. The most apparent contrast exists between Community 1, which heavily features UK news sources, and Communities 2 and 3, which include mostly US sources. The distribution of words and bigrams in Figs. 5.6b and 5.6c show some topical differences between the four communities in the left-wing cluster. Communities 3 and 4 include more terms related to the consequences of the decision for the *American* people whereas Community 1 is mainly concerned with the international political ramifications. These findings suggest greater focus and coherence amongst the right-wing and climate-sceptic frames and support the findings of [70], with greater fragmentation amongst the left-wing and environmentalist frames; this may partly reflect the larger size of the left-wing/environmentalist cluster in the information-sharing network, which permits greater internal differentiation.

5.3.3 Consistency of network structure over time

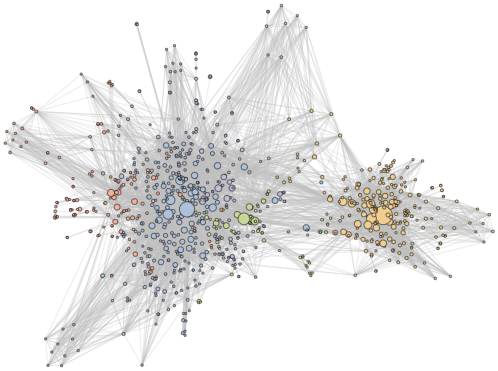
To understand the consistency of the polarised information-sharing process over time, the network structure was examined along with similarity/persistence of the sets of users, source domains and shared URLs (articles) across all weeks in the study period. Similarity scores are reported in Table 5.3. Fig. 5.7 presents the network diagrams of the remaining six weeks in our study period (Weeks $1-3$ & $5-7$).

In each week, a similar division into two clusters is observed in the information-sharing network, although the specific composition of the users and shared articles that form the network changes substantially over time. The number of communities in the left-hand and right-hand clusters varies over the weeks and the relative sizes of the different communities also change. However, each week reveals the same broad pattern of a larger left-wing/environmentalist cluster split into several smaller sub-communities, with a smaller right-wing/sceptic cluster, showing that this pattern is not an artefact of the increased activity during Week 4. Taken together with Fig. 5.3, these results demonstrate that the pattern of network division persists over

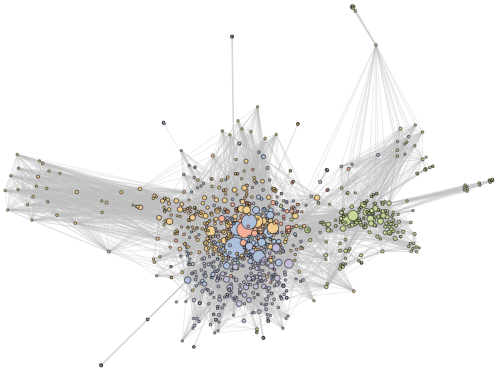
time, spanning multiple weeks of 'normal' activity and one exceptional week of high media activity.

Considering the inter-week similarity between user populations and the source domains and URLs they shared, Table 5.3 shows the inter-week similarity of unique users, URLs and website domain shares, along with the minimum, mean, maximum and standard deviation of pairwise similarity measures. The shared articles in Table 5.3b show the lowest similarity between weeks, with a slightly higher similarity between adjacent weeks. Intuitively, the source domains from which content was shared show higher similarity between weeks in Table 5.3c. User populations show limited persistence between weeks, summarised by Table 5.3a. The overall pattern is that, from week to week, a limited proportion of the user population and set of source domains persist, with much lower persistence in the sets of articles that are shared. This is an interesting finding with respect to the high consistency in the two-cluster network structure that is reliably observed every week. The lack of persistence in shared URLs may be explained by the general volatility of news media, where articles typically have a short lifetime (e.g. 2-3 days visibility in online sharing [187]). Moderate persistence of users and sources between weeks perhaps suggests an active core group who are present each week, with a wider group who appear less frequently.

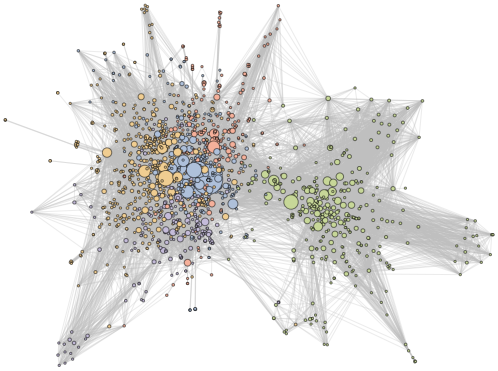
A marked difference can be observed in the typical similarity scores for Week 4 and for other weeks. Lower similarity scores were observed for users, URLs and domains from Week 4 appearing in other weeks, while conversely, higher similarity was observed for users, URLs and domains from the other six weeks appearing in Week 4. The number of new users, URLs and domains in Week 4 also shows a stark contrast with other weeks: 70.1% of users, 82.0% of URLs and 47.9% of domains are unique to Week 4 and not present in any other week. This is strong evidence that the events of Week 4 appear to have spurred an influx of both social media participants and digital media sources (users: mean 49.3%, min. 45.6%, max. 51.2%, $\sigma = 2.10\%$; URLs: mean 70.3%, min. 66.6%, max. 73.2%, $\sigma = 6.54\%$; domains: mean 29.8%, min. 27.3%, max. 31.5%, $\sigma = 4.21\%$). While web domains exhibit the



(a) Week 1: 60.24% of 1,660 nodes visible.

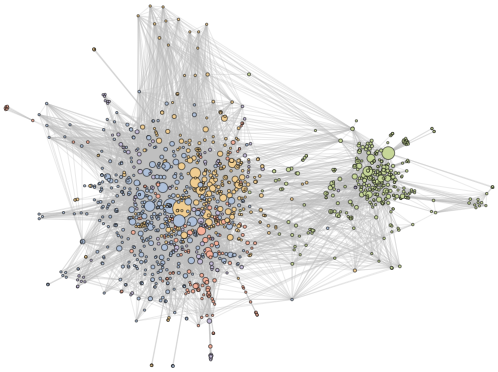


(b) Week 2: 66.04% of 1,802 nodes visible.

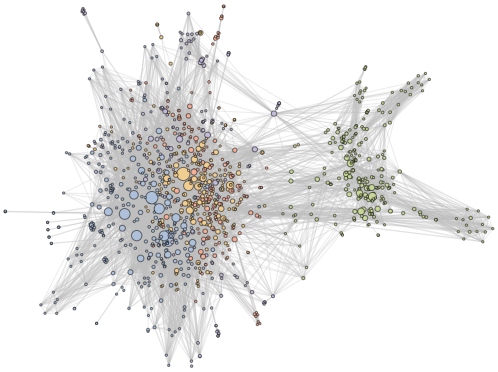


(c) Week 3: 67.40% of 2,362 nodes visible.

(d) Week 5: 67.71% of 2,205 nodes visible.



(e) Week 6: 67.07% of 2,092 nodes visible.



(f) Week 7: 66.16% of 1,856 nodes visible.

Figure 5.7: **Consistent network structure over time.** Network diagrams of the top five communities across the six remaining weeks. Each figure is oriented such that the left-wing cluster is on the left and the right-wing cluster is on the right. In each case node colour signifies community membership and size is proportional to the square root of total share count. Communities are labelled $1-5$ in decreasing order of size, and coloured blue, yellow, green, red and purple respectively. Node placement is determined by the Python implementation of the ForceAtlas2 algorithm [92].

most stability between weeks, URLs exhibit the least stability and users fall between these two extremes.

To confirm that the persistent two-cluster network structures seen in Fig. 5.7 are polarised along the same left-right political and environmentalist-sceptic ideological axes as that seen for Week 4 (see Fig. 5.5), the domain bias codings were applied to the networks created for Weeks 1-3 and 5-7 (Figs. A.6 and A.7). Furthermore, average biases were calculated for all communities in each week in Fig. 5.8. Trends in polarisation over time are shown as network averages for political and climate change biases, alongside the community averages, in Fig. 5.8. In most weeks, the whole-network average shows mild left-wing and environmentalist bias. Week 5 shows (for the only time across the seven-week period) a large neutral community. Supporting the visual evidence seen in Figs. A.6 and A.7, and the community-level bias scores in Fig. 5.8, these findings show that the polarised network structure observed in Week 4 is persistent and is not an artefact of the increase in activity in the climate change conversation despite the turnover in users, source domains and shared articles.

5.4 Discussion

This paper presents an analysis of digital media sharing behaviour around the contested issue of climate change. Our analysis looks at the network structure formed by users sharing web articles related to climate change, combining network analysis with computational and human text analyses to identify and characterise communities of users and the articles/sources they share. The aim is to understand how people engage with, and share online media content about climate change on social media (specifically Twitter).

We have found that amongst the communities of shared URLs, right-wing and climate sceptic views are strongly correlated, as are left-wing and environmentalist views. This correlation has been observed in individuals before using survey-based methods (e.g. [54, 64, 88]). Our study shows that the association extends to media outlets, specifically to the content produced by a large set of online news providers

Week	1	2	3	4	5	6	7
1		0.29	0.33	0.39	0.25	0.25	0.22
2	0.24		0.30	0.35	0.22	0.24	0.21
3	0.19	0.21		0.38	0.20	0.19	0.18
4	0.08	0.09	0.15		0.11	0.10	0.09
5	0.18	0.20	0.25	0.36		0.24	0.21
6	0.19	0.22	0.26	0.35	0.26		0.25
7	0.20	0.23	0.28	0.36	0.27	0.30	

$\min.(S) = 0.084$
$\text{mean}.(S) = 0.236$
$\max.(S) = 0.389$
$\text{stdev}.(S) = 0.078$

(a) Asymmetric overlap of users.

Week	1	2	3	4	5	6	7
1		0.13	0.08	0.11	0.06	0.05	0.04
2	0.10		0.14	0.10	0.05	0.05	0.04
3	0.05	0.10		0.16	0.05	0.04	0.03
4	0.02	0.03	0.06		0.06	0.03	0.02
5	0.04	0.05	0.06	0.19		0.16	0.07
6	0.04	0.05	0.06	0.10	0.17		0.16
7	0.04	0.04	0.05	0.09	0.09	0.19	

$\min.(S) = 0.022$
$\text{mean}.(S) = 0.078$
$\max.(S) = 0.193$
$\text{stdev}.(S) = 0.047$

(b) Asymmetric overlap of URLs.

Week	1	2	3	4	5	6	7
1		0.55	0.59	0.72	0.58	0.54	0.51
2	0.44		0.58	0.75	0.48	0.55	0.48
3	0.34	0.41		0.76	0.42	0.42	0.39
4	0.15	0.20	0.29		0.23	0.22	0.18
5	0.41	0.43	0.53	0.76		0.52	0.46
6	0.41	0.52	0.56	0.76	0.56		0.53
7	0.46	0.54	0.62	0.77	0.59	0.63	

$\min.(S) = 0.155$
$\text{mean}.(S) = 0.496$
$\max.(S) = 0.767$
$\text{stdev}.(S) = 0.158$

(c) Asymmetric overlap of domains.

Table 5.3: Similarity scores for users, URLs and domains between each of the seven weeks calculated using Equation 5.3. Similarity is directional. The similarity given in cell $(4, 1)$ is the proportion of the users/URLs/domains in Week 4 also in seen Week 1. Cool shades indicate values smaller than the mean while warm shades indicate values greater than the mean.

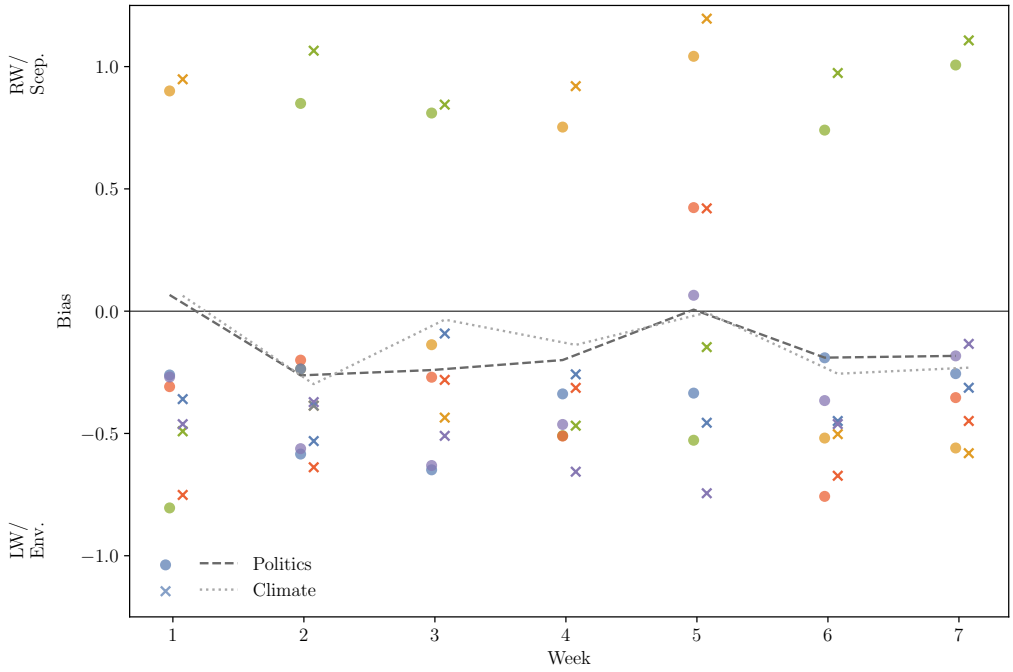


Figure 5.8: **Levels of political and climate change bias over the course of the seven week study period.** These are measured as the mean coded bias of domains weighted by total shares (see Section 5.2.6). Bias in each of the five largest communities are represented by the scatter points in each week, and the bias across the whole network is given by the lines. The colour of the community points is consistent with other figures. In most weeks, the average network bias is left of centre and more environmentalist than sceptic. This trend continues to the individual communities, with the majority being left-wing and environmentalist.

(Fig. 5.4), and is a major feature of how online content is shared. Recent work by Hornsey et al. [88] suggests that the scepticism-conservatism link may be strongest in the US. This finding is supported here, where we observe two large communities based mainly on US sources; one more left-wing and environmentalist, the other more right-wing and sceptic. Although we did not examine where the users in our dataset were located geographically, the dominance of US sources in two of the communities, and of UK sources in one of the communities, suggests that these communities are largely formed around users from those countries. Our overall focus on English-language tweets and media content, plus large numbers of Twitter users in these countries (nearly 15% of online US adults used Twitter in 2017 [167]), are consistent with this conjecture.

In this study we find two levels of community structure. At the highest level, there is polarisation and segregation with a large left-wing/environmentalist grouping and a smaller right-wing/sceptic grouping clearly visible in the sharing network after the application of the force-directed layout algorithm. Within the left-wing/environmentalist group, algorithmic community detection finds (typically) four smaller groupings, characterised by further analysis as above. It is interesting to note that the right-wing/climate sceptic group is more densely connected internally than the left-wing/environmentalist group (Student's t-test applied to average clustering coefficients of all left-wing and right-wing communities in Figs. 5.3 and 5.7, with clustering coefficients calculated in Gephi using Latapy's algorithm [106], $p = 0.019$). This finding echoes the observation by Colleoni et al. [43] that in the US, active Republicans exhibit greater levels of homophily than Democrats in their patterns of interaction on Twitter.

The temporal analysis presented above gives confidence that the polarised network structure is robust over time, despite turnover in the user populations, sets of shared articles and news sources. We studied similarity between the sets of users, URLs and domains across the seven weeks, as well as the proportion of users, URLs and domains appearing only in the exceptional Week 4 (in which climate change was the subject of a mainstream news event when President Trump announced the US

withdrawal from the Paris Agreement). This major event substantially increased the volume of social media messages related to climate change and drew in an increased number of new users, who shared a greater number of articles from a greater number of online information sources. Thus we observed network structure and composition in both 'normal operation' as well as in an unusual state of high activity, showing the typical level of week-to-week volatility as well as the substantial change under the influence of the disruptive mainstream news event. At all times, including the disruptive event, the polarised network structure remains strong and clearly visible, despite substantial changes in its constituent parts (users and articles). As such, we conclude that the broad topological features visible in Figs. 5.3 and 5.7 are stable over the seven-week duration of our dataset. Similarly, we find that the association of left-wing/environmentalist views, and of right-wing/sceptic views, are persistent, as shown in Figs. 5.5a, 5.5b, A.6 and A.7. Future study may find it fruitful to examine whether such stability exists for other online networks or other politicised discussions.

The similarity statistics in Table 5.3 suggest a general trend for a small proportion of the users, URLs and domains to persist across weeks while others appear only sporadically. In the URL case, the comparatively low levels of similarity are not surprising as news articles are quickly superseded by new facts and perspectives. Among the users and domains, the evidence suggests the network has a stable core with an unstable periphery. Over the seven weeks studied, a number of websites establish themselves as critical to the flow of information around climate change, either as news sources (such as *The Guardian* or *Breitbart*) or as conduits for personal opinions (such as *Twitter* and *Wordpress*). Many other sites are peripheral to the news-sharing network of climate change, appearing sporadically with a lower frequency of usage. This paper has deliberately avoided studying users as individuals, but it seems reasonable to expect a core group of committed and strongly interested people who regularly share information about climate change, with others who contribute less frequently or only when motivated by external factors. Considering this user behaviour in the context of the persistent network structure we have observed,

it appears that the revealed polarisation is a feature of the system as a whole and not caused by specific events or users.

This article illuminates several new dimensions of the media debate around environmental politics and climate change. The findings complement previous studies that have shown climate-related echo chambers exist in direct user-user interactions (e.g. [50, 196]) by showing that similar structures also characterise patterns of information-sharing. The Twitter messages that form the sharing network studied here reflect a mechanism of news promotion and active attempts to inform others, which is not always the case in direct personal interactions. We also capture a vital additional element by exploring the network topology over time and demonstrating that the polarised structure persists, even when there is a disruptive mainstream news event. In addition, we have shown a strong correlation between political and climate-related ideological biases in news production and consumption, with associations between left-wing/environmental and right-wing/sceptic positions. This widens the scope of previous studies of mainstream news media (e.g. [64]) to include the increasingly important online news media.

When looked at in isolation, several domains appear to be coded with more of a left-wing, environmentalist bias than would be expected by domain knowledge experts. Typically neutral sites, such as Bloomberg, are reported as left-wing and environmentalist primarily through the lenses they use to cover the specific topic of climate change. Other domains, such as the Daily Caller, may have been poorly represented by the content sampling process as factual and scientific extracts taken without surrounding challenging context were considered as supporting the scientific consensus. This "good faith" on the part of the coders returns a particularly erroneous response for Watts Up With That, a prominent blog in the sceptical community. Many of their articles present quotations from public figures and scientific papers in addition to their own commentary that frequently challenges the framing, which are contextual clues that may be unavailable to the coders. The last explanation for possible misclassification of certain domains identified in this exercise is the reposting of content from other sources. Investigation of the articles coded from Fox

News found that three of the five articles were from Associated Press sources, which may present a different editorial bias to Fox News original content. The choice of which content to repost from other sources is an important editorial decision and as such we believe this inclusion has not adversely affected our findings.

Future work in this area could further examine the ties between political and environmental opinions. This study used human coders to detect biases from article content. While this approach was successful, it carries substantial costs which make it hard to operate for large datasets. It could be argued that our use of subjective human grading is a limitation in our analysis, but it is a necessary compromise given the inherently subjective nature of political and environmental beliefs and the current lack of objective tools for analysing such complex interpretations in large quantities of data. An automated classifier for these biases would require significant work for its creation (using a blend of machine learning and natural language processing), but would support future large-scale studies, a necessary step given the ever-increasing volume of online content. Another research question concerns whether the patterns observed on Twitter extend to other online platforms. We studied Twitter due to its prominent role as a means for users to find news stories [168] and its frequent usage as a platform for lively political debates in which strangers can interact. If suitable datasets could be obtained, it would be possible to apply similar methods to other popular sites such as Facebook and Reddit; however, privacy restrictions prevent research on many social media platforms. Finally, one further question that is not addressed here is downstream exposure to shared content, that is, who sees the articles that are shared by Twitter users? It is probable that the users captured in our dataset, that select and share online articles about climate change through Twitter, act as 'opinion leaders' (following the long-established two-step flow model of Katz and Lazarsfeld [96]), locating new information about the topic and disseminating it to their followers. Measuring the volume of downstream views and retweet rates would answer important questions about which messages are more effective when shared on social media. One key question is the extent to which polarisation exists in this secondary consumption of climate change media, or whether network

effects mitigate polarisation by capturing views from both sides.

Any gathering of data from online social networks requires careful consideration of the biases that it may introduce. The users of online platforms are a different distribution of people than the general population: they tend to be younger, more wealthy and better educated [198]. Our sampling on keywords means that we will capture users who are more engaged with the topic of climate change, and moreover are significantly invested enough to share information with their followers. As such we are considering a highly-motivated sample of a certain part of society, but this is what will be visible to many users when they visit Twitter. This sampling also explains the difference from Global Warming's Six Americas [112] that we observe, as only the "alarmed" and "dismissive" types are likely to be invested enough to be captured. The work of Williams et al. [196] support this as their coding of Twitter users found only the extreme environmentalists and sceptics. Our keywords were chosen to capture as much of the English language components of the climate change conversation online, but the observed structures may vary in other languages. These factors should be considered in any analysis of social media data but we do not believe that they affect the strength of our findings in any way.

This work demonstrates that media communication of information around climate change faces many challenges in the age of social media. Users return to the same trusted sources for information even when presented with new contexts, and any attempts to attract new readerships need to consider this behaviour. Understanding how these trusted ties form will be key to combating the spread of misinformation that currently challenges online social networks and vital for promoting action on climate change and other societal challenges.

Short- and long-term collective attention

What does the previous chapter tell us about collective attention?

In the previous chapter, we explored the relationship between Twitter users and the URLs they share in more detail. Using a period of increased attention to the climate change debate as a case study, we have seen how the typical structure and polarisation is affected by a disruptive media event.

In particular, we have seen that President Trump's announcement of his intention to withdraw the United States from the Paris Agreement on climate change lead to an approximately ten-fold increase in the amount of tweets per day containing the strings *climate change* or *global warming* (Fig. A.1). Alongside this increase in engagement with the climate change debate, we track the polarisation in sources of information and their position in the network. We have seen that away from the week of the announcement (when activity was at its usual baseline level) source domains divide themselves into two spheres: a left-wing, environmentalist sphere and a right-wing, sceptic sphere. These two spheres are evidence that users are likely to engage with one side of the climate change bias spectrum almost exclusively, with comparatively few users bridging the opinion divide. This pattern persists across the seven weeks studied and is robust under the perturbation induced by the ten-fold increase in tweet activity (Figs. 5.5, A.6 and A.7).

Given the temporal networks available in this dataset we can also answer ques-

tions about how the increased activity affects the participants in the network (in terms of both users and URLs). We observed a high turnover in the URLs shared in each week (Table 5.3), which is unsurprising given how rapidly the news cycle updates. We saw evidence that Week 4, and its associated increase in media awareness to climate change, involved a higher rate of return for users and the source domains they share compared to any other pair of weeks. Most importantly of all, Week 4 saw a majority of new participants and URLs, and almost 50% of domains did not appear in any other week. These new participants in the conversation were shown to fit within the existing polarised structure, and any impact to reduce polarisation was restricted to areas away from the largest communities.

This short period of heightened awareness improves our understanding of collective attention more generally. We have shown that while increased collective attention draws more participants into the debate, some network structures are sufficiently robust that this influx does little to disrupt them. In addition, we observe differences in how attention to different features manifests. Certain topics and terminology were shown to be inherently linked to positions in either political or climate change bias. This improves our understanding of framing in such politicised debates and shows that attempts to model attention in such systems must account for the selection bias of individuals.

Through this study we have seen that the disruption to the normal conversation patterns around climate change was short-lived, despite the large absolute increase in activity. It remains to be seen how social media interactions and collective attention respond to longer-term stimulus.

How does this inform our upcoming work?

The COVID-19 pandemic of 2020 took many members of society by surprise. For many in Europe, the first sign that widespread restrictions may be required was the lockdown of several regions of northern Italy. These restrictions quickly spread to cover all of Italy in early March 2020, and similar measures were soon in place across the continent. Varying levels of restrictions have been in place around the world

for the rest of 2020 and into 2021, and therefore present the perfect opportunity to study the longer-term impact on attention of a disruptive media event.

In the next chapter we seize this opportunity to study how collective attention on social media responds over longer timescales. We will measure the impacts of this disruption on attention in several ways, using England and Wales as a case study. Firstly, we look at the patterns of tweet locations to understand how mobility restrictions affect attention to different locations tagged in social media content. To capture changes in mobility we consider the trajectories of individual users to find out how far, and fast, they travel between place tags as a means of using social media data to quantify the impacts of social distancing restrictions. The next chapter will also explore the change in collective attention to different topics by constructing networks of hashtag usage over time. Considering these three aspects of social media data will improve understanding of how the ongoing pandemic, changes in the severity of lockdown efforts, and the associated news cycle, are manifested in the choice of conversation topics around high-impact media events over timescales longer than the typical days or weeks.

Chapter 6

A tale of two lockdowns: Comparing mobility and attention responses in two periods of coronavirus restrictions in England and Wales

Abstract

In order to combat the spread of the COVID-19 pandemic across the globe, restrictions have been put in place to limit the social contacts of individuals. As a result, many forms of social interaction were forced to take place online. Through the study of a dataset of geolocated tweets in England and Wales we investigate how the patterns of mobility seen in Twitter are affected by the different levels of lockdown restrictions in place from 26th March 2020 until the second full lockdown in November 2020. We find that lockdown restrictions decreased the visitation in many areas in the UK, and particularly rural areas or holiday destinations, with limited effect in periods of weakened restrictions. This result is supported through measuring travel between successive tweets which show limited change in the distances travelled by individuals in the autumn lockdown. Using hashtags to capture topics of interest

we show that this behaviour may arise from a transition of the coronavirus from an exceptional event with heightened awareness early in the pandemic towards common circumstances as the situation persists.

6.1 Introduction

The early months of 2020 quickly demonstrated that the coming year would see significant disruptions to modern life. First reported in Wuhan, China in January 2020 the COVID-19 pandemic soon spread to affect countries across the world. Efforts to limit the spread and protect public health varied from country to country, but the nature of the responses typically followed changes in knowledge of the infection and current case numbers. Most countries introduced some form of social distancing measures and self-quarantine expectations to limit opportunities for contact between citizens and thereby limit new infections. These restrictions often came in the form of "stay at home" orders, where individuals are asked to only leave their homes for essential reasons such as medical appointments and to work from home wherever possible.

The various social distancing restrictions will have had an impact on the mobility trends of citizens on both local and global scales. Measuring this reduction gives an additional indicator of the success of these efforts beyond reductions in case numbers captured through traditional testing regimes. Furthermore, understanding how different levels of social restriction have affected mobility can give insights into the level of adherence to the restrictions, especially given the effects of misinformation on the course of the pandemic [153]. Beyond consideration of mobility, the scale of the pandemic is unknown in recent memory and therefore has disrupted the news cycle. Understanding how attention to the continued news updates varies over the course of the pandemic adds a valuable perspective into how public health messaging is received. This manuscript will explore these questions through the use a large social media dataset.

6.1.1 Background

Measuring human mobility presents several challenges to scholars. One primary difficulty is the availability of suitable location data. Typically, call records of mobile phones are used to infer locations from nearby towers. Although these data require agreements with the network providers (and are therefore not widely publicly available), the location data within have been shown to capture underlying spatial patterns. Expert et al. [59] found that the network structure of call interactions recovered the split between the French and Dutch speaking parts of Belgium. Other uses of mobile phone data include tracking mobility around cities (e.g. [73, 208]), but at this scale concern for the privacy of individuals must be taken into account. Through aggregation methods, this concern can be limited, but cannot be completely eliminated when there is a possibility of function creep in the applications of such data [176]. Other data sources include traveller numbers in various transport networks (e.g. airline travel [16]), although such attempts are limited in their difficulty to capture short-range mobility using individual vehicles, cycling or walking.

Social media presents one potential avenue to provide higher density location data. Platforms such as Twitter, Facebook, Instagram and Foursquare all allow users to provide various levels of location data with their posts. Foursquare users choose to check-in to various locations, and through this action movement from location to location can be tracked. Previous work has shown that these mobility patterns can reveal spatially coherent relationships between locations [116]. This data arises from an underlying social network; reconstructing this network can help with linking observed patterns in the social data to similar patterns known in offline scenarios [7]. Some work has looked at using location data from Facebook (e.g. [60]) and Instagram (e.g. [207]) but these are comparatively rarely used sources given the increased level of privacy expected by users and the increased difficulty of accessing data. Twitter locations are the most commonly studied (e.g. [129]) due to the open nature of the platform and the ease with which data can be gathered. Much work has shown that the mobility patterns recovered through Twitter are representative of those seen in other media (e.g. [83, 95]), but studies also highlight challenges with such analysis.

In particular, Armstrong et al. [12] identify limitations in the use of Twitter data to measure migration patterns. The difficulties stem from a number of individuals being incorrectly identified as migrants based on their geolocation histories. The authors rightly question whether this issue arises from collection biases or the irregularity of posts by most users, but improvements could also be made in the algorithms used to identify migrants. Beyond mobility, location-based social media data has been used to develop "social-sensing" tools which aim to augment traditional sensor networks for tracking and measuring the impacts of natural disasters (e.g. earthquakes [126] and storms [170]).

The COVID-19 pandemic is not the first time that data sources for mobility patterns have been used to study disease dynamics. Wesolowski et al. [195] give a good overview of how mobile phone data can be used to improve the modelling of infectious diseases, noting that despite the potential bias in the individuals that own mobile phones, advantages from the finer temporal and spatial scales can prove invaluable for studies at the population level. More recently, mobile phone data has been shown to demonstrate that travel restrictions in Sierra Leone to combat the spread of the ebola virus did reduce mobility [147]. Twitter has been well-established for capturing trends in pandemics such as estimating flu cases [103], hayfever symptoms [48], dengue fever [76] and HIV [204].

Given the global scope of the COVID-19 pandemic, a number of studies have already explored different aspects of societal responses to the infection. In addition to tracking the cases and symptoms, location data have been used to provide additional context to the course of the pandemic such as how different groups of society respond to social distancing restrictions (e.g. [39, 163, 191]). Beyond case tracking efforts, the rich content of tweets has enabled efforts to measure how public opinions change over time. Dyer and Kolic [55] track the linguistic tones used by individuals and track the transition from emotional phrasing towards analytical terms as the pandemic continues. The more negative aspects of social media exposed through the spread of misinformation have also seen study. One notable area is the coexisting pandemic of misinformation. In particular, scholars have explore how misinforma-

tion can limit the effectiveness of preventative measures [153] and how the social structures aiding diffusion can be disrupted [203].

6.1.2 Motivation and research questions

There remain understudied aspects of how social distancing restrictions affect mobility as seen on social media, and how the changes are sustained over longer periods. Throughout the course of 2020, the level of coronavirus restrictions in place in England and Wales fluctuated as case numbers rose and fell. After increasing closures of certain businesses and localised advice, a coordinated lockdown was introduced across both nations on 26th March 2020 requiring citizens to limit their time outside the home to essential trips only. These restrictions were in place until the summer, before being gradually relaxed as levels of infection in the community declined. The national picture of coronavirus had changed by the autumn, with some divergence of regulations between the two nations. This period also saw increasing numbers of coronavirus cases necessitating the reintroduction of restrictions. Restrictions applied at the local level became the preferred approach in September and October until rising case numbers forced another period of national lockdown beginning on 23rd October 2020 in Wales and on 5th November 2020 in England.

In this manuscript we will explore how the coronavirus pandemic and subsequent social distancing restrictions have affected social behaviour on Twitter. To derive some insight into how society has been impacted by the pandemic we study the data through three lenses: spatial attention, mobility and topic attention. With the requirement for most people to stay at home as much as possible early in the pandemic, we would expect to see some changes in which locations users choose to tag in their tweets. In particular, we should find that some areas see a change in the relative number of tweets compared to an appropriate period of normal activity. This assumption motivates our first research question.

- RQ1: How does attention to different locations on Twitter change in periods of lockdown?

Beyond studying where individuals tweet from, we can also use these locations

to measure how individuals move around England and Wales. As with location attention, we would expect to see the lockdown restrictions leading to fewer and shorter journeys by the majority of individuals who should be staying at home. However since the offline lockdown restrictions do not directly affect online tweeting behaviours, the strength of this association is unclear. We will address this situation in our second research question.

- RQ2: How does mobility of Twitter users change in periods of lockdown?

To complement the spatial attention and mobility components covered in the first two research questions, we make use of the rich topic data available in Twitter hashtags to track how the lockdown events alter the visibility of certain topics. We build on the ideas of Borge-Holthoefer et al. [26] and consider the structural changes in the network of individuals and the hashtags they use over the course of major events in the news cycle. Understanding this topic behaviour motivates our final research question.

- RQ3: How does attention to key topics in the coronavirus pandemic change between two lockdown periods?

The rest of the manuscript proceeds as follows. In Section 6.2 we outline our data source and detail the methods used to track user locations, trajectories and the aggregate trends in attention during the lockdown periods. In Section 6.3 we report the findings in each of these three areas, before discussing the insights we can draw from these results in Section 6.4.

6.2 Methods

6.2.1 Dataset and preprocessing

This work uses a dataset of tweets with location data, primarily from England and Wales. We used a location filter on the Twitter Streaming API to capture all tweets located within the bounding box of with bottom left longitude and latitude

(-5.8,49.9) and top right longitude and latitude (1.8,56) for four three-month periods over 2019 and 2020. These periods cover April-June and September-November in each year. We also collect tweets from March 2020 to track the emergence of the COVID-19 pandemic as a major event and the introduction of the first restrictions to combat its spread. The collection was interrupted for most of the final 10 days of September 2019 and a period of two or three days in each of April, May, June, and November 2020. It is not anticipated that this will invalidate any conclusions drawn from the data since the remaining data in these periods includes several million tweets and there are no patterns in which times the collections failed (i.e. no systematic bias is expected in the resulting dataset).

6.2.2 Location inference

Location data in tweets is provided by users in one of two ways. If the user allows such data to be collected, a GPS tag based on device location is included in the tweet metadata. This information is included in only a small proportion of tweets (typically 1% in datasets with no location filter [170], and 6.2% of tweets in this dataset). The more common means with which tweets include location data is through the use of 'place tags'. When a user chooses to provide this information, they are presented with a list of place tags defined on Twitter. Each of these tags correspond to a (usually) rectangular bounding box that contains the named location. These bounding boxes can be of various sizes, and as such can lead to varying accuracy in any subsequent point estimates.

We handle these two different types of location data in different ways to maximise location accuracy while maintaining a large sample size. In order to ensure a minimum level of accuracy in the place tags, we reject those location tags whose bounding box diagonal is greater than 30km. This limit was chosen to retain most of the place tagged tweets, while removing the least accurate locations in light of the distribution of bounding box sizes found by Weaver et al. [189]. In our dataset, this limit retained 91.5% of place tags. Through this step we remove coarse level locations (typically at the county or country level in the context of the UK) but still

retain locations at the level of smaller regions and large cities like London. Each bounding box is then converted to a single point by averaging each of its vertices. In most cases, the bounding box is a rectangle and therefore we assign the tweet to the centroid of its bounding box. Since tweets with GPS locations give only a single point, no additional preprocessing is required. After this filtering step, we are left with 61,030,475 tweets.

6.2.3 Location and trajectory analysis

Now that we have converted the bounding boxes of place tags to coordinate points, we are in a position to explore how patterns of offline mobility are reflected online, and in particular how the different levels of lockdown restrictions affected these patterns in comparison to a normal year. We first measure the spatial distribution of user locations within the England and Wales bounding box used for data collection. The collection area is divided into a grid of 50x50 cells, with the total number of tweets appearing in each cell counted in a given month. This resolution is such that the diagonal length of each cell is between 16 and 18km, which is greater than the maximum possible error in tweets using place tags. When comparing between years we normalise the counts to account for changes in the overall level of Twitter activity. By dividing the count in each cell by the total number of tweets for the month we are left with a measure tracking the proportion of observed tweets appearing in each cell.

To measure mobility, we compile sequences of tweets during the course of each calendar month, each linking successive tweets by a single user. This is done after all other preprocessing steps, to exclude any tweets without sufficient location accuracy. Using the between-tweet intervals we can calculate the elapsed time and distance between each pair of successive locations, and thereby estimate the speed the user has travelled. The calculated values for distance and speed are necessarily lower bounds for the (unknown) actual values, since these calculations assume travel along the shortest path at a constant rate.

Calculating these values over the tweet sequences of all users in the dataset al-

lows us to move beyond looking at the spatial distributions of users alone. The population-level characteristics for the distances travelled and the speeds at which they do so contribute an additional component to the comparison of pandemic conditions to normal mobility and address how journey types are affected by lockdown restrictions.

6.2.4 Discussion networks

Our final perspective for understanding how the COVID-19 pandemic and ensuing public health measures affect attention is through a network representation of conversation topics. We use a bipartite network representation of individuals and the hashtags they include in their tweets. In this case, we choose hashtags as the indication of different topics over tokens extracted from the tweet text as hashtags are conscious indications that these terms are important to the message of the tweet. We construct these networks at the day level to better reflect the shorter timescales of the news cycle.

Given these networks we are interested in whether there are groups of users who share common topics of interest. We approach this from the direction of modularity-based community detection, a method of finding densely connected subsets of the network using no knowledge beyond the network topology. Bipartite networks introduce an extra complication to this approach due to their additional requirement that edges between two users or two hashtags are not allowed. To account for this restriction, we use Barber's modularity [21] which has been designed to respect the structural constraints of bipartite networks. Using this method, we can identify groups of hashtags which are more likely to be used by similar groups of users, and compare these structures over time. The communities detected contain both user and hashtag nodes, but in order to avoid focusing on users by name we only consider the hashtags for any node level analyses. Modularity-maximising community detection will separate components in the network into distinct communities and to avoid this affecting the analyses we consider only the giant component for the community calculations.

In order to analyse the networks and their community structure we make use of several additional metrics. To enable some understanding of the relative sizes of each community in the network we calculate the expected community size, given by

$$\frac{\sum_C |C|^2}{(\sum_C |C|)^2}. \tag{6.1}$$

The expected community size ranges from 0 to 1 and indicates the expected proportion of all nodes in the community of a randomly selected node. To measure the general level of interest in coronavirus on a given day, we check whether the hashtags used contain any of the following terms: *coronavirus*, *covid*, *lockdown*, *pandemic* and *epidemic*. This list is not exhaustive but was chosen to capture key themes with limited applicability in off topic contexts.

6.3 Results

6.3.1 Location patterns

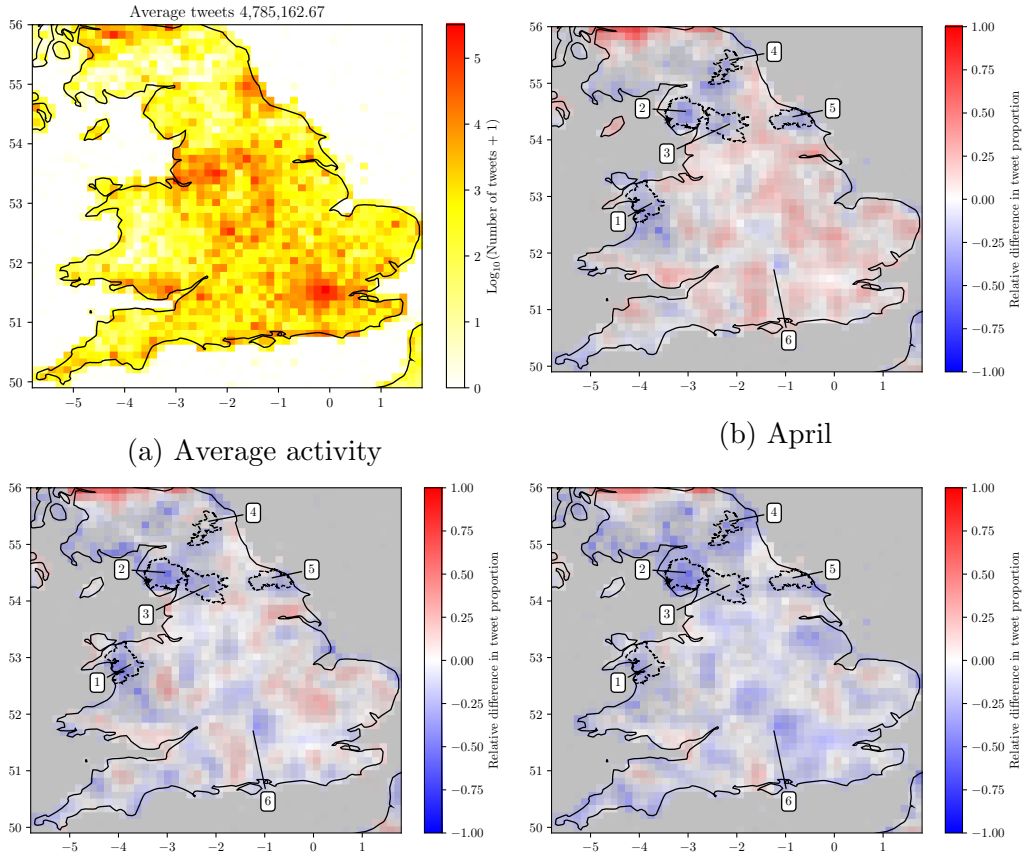
We begin by exploring how the restrictions imposed to combat the spread of COVID-19 in England and Wales have affected the distribution of locations on Twitter. Under normal circumstances, we would expect the number of tweets from a location to be closely linked to the population density in the area.

We measure Twitter location activity by dividing the collection bounding box into a rectangular grid of 50x50 cells. Using the distribution of user locations across this grid, we compare the months of April, May and June in 2020 to the same months in 2019. During this period in 2020, England and Wales were placed under a series of social restrictions that required individuals to stay home and limit their contact with others outside their household. Fig. 6.1 shows the comparison of this period over two years of Twitter activity, with a Gaussian smoother applied to reduce local noise. We first notice from Fig. 6.1a that the distribution of Twitter activity closely follows population density. Considering Figs. 6.1b, 6.1c and 6.1d we see an evolving pattern in how the attention to different locations has changed between the two years. As the lockdown progresses, we see that the areas of increased activity in 2020 (red)

initially cover much of the south-east of England, and typically the activity shows smaller changes where the underlying population is largest. One exception to this trend is Oxford (point 6), possibly due to the early departure of students before national restrictions were in place[1]. Over the course of the next two months, the lower activity first seen in the national park areas of Snowdonia (area 1), the Lake District (area 2), the Yorkshire Dales (area 3) and Northumberland (area 4) extends across much of England and Wales. The patterns in these initial locations likely arise from a lower baseline population and the limitations on recreational trips imposed by the lockdown restrictions. Through this we can see that initially the change in location attention is determined partially by the typical reason for visiting, and this link declines as the pandemic continues.

After a summer of relatively relaxed restrictions across most of the country, increasing levels of restrictions emerged in England and Wales throughout the autumn, culminating in a full lockdown for four weeks starting on 23rd October in Wales and 5th November 2020 in England. The difference in Twitter activity between 2020 and 2019 for these months is shown in Fig. 6.2. As we see in Fig. 6.2a, the average activity level during this three month period follows the same pattern as April, May and June by closely correlating to the densely populated urban areas. The change in location attention is different during this period compared to the first lockdown. Initially much of Wales and southern England sees less attention in 2020 compared to 2019. Of note is the increased attention to national parks in northern England. In contrast to April, September sees increased visitation to the national park areas of the Lake District (area 2), Northumberland (area 4) and the North York Moors (area 5). For November, during which time most of the country was under stay-at-home orders, Fig. 6.2d shows a similar pattern of increased attention in south-east England as that seen in Fig. 6.1b. This is in stark contrast to Figs. 6.2b and 6.2c which show that these regions saw less activity in 2020 than in 2019 during these months. One final pattern that can be observed in Fig. 6.2d is the generally lower colour intensity over much of England and Wales. This suggests that the relative change in tweets from these locations is smaller compared to the other months we

[1]See communications for 13/3/20: `https://www.ox.ac.uk/coronavirus/communications`



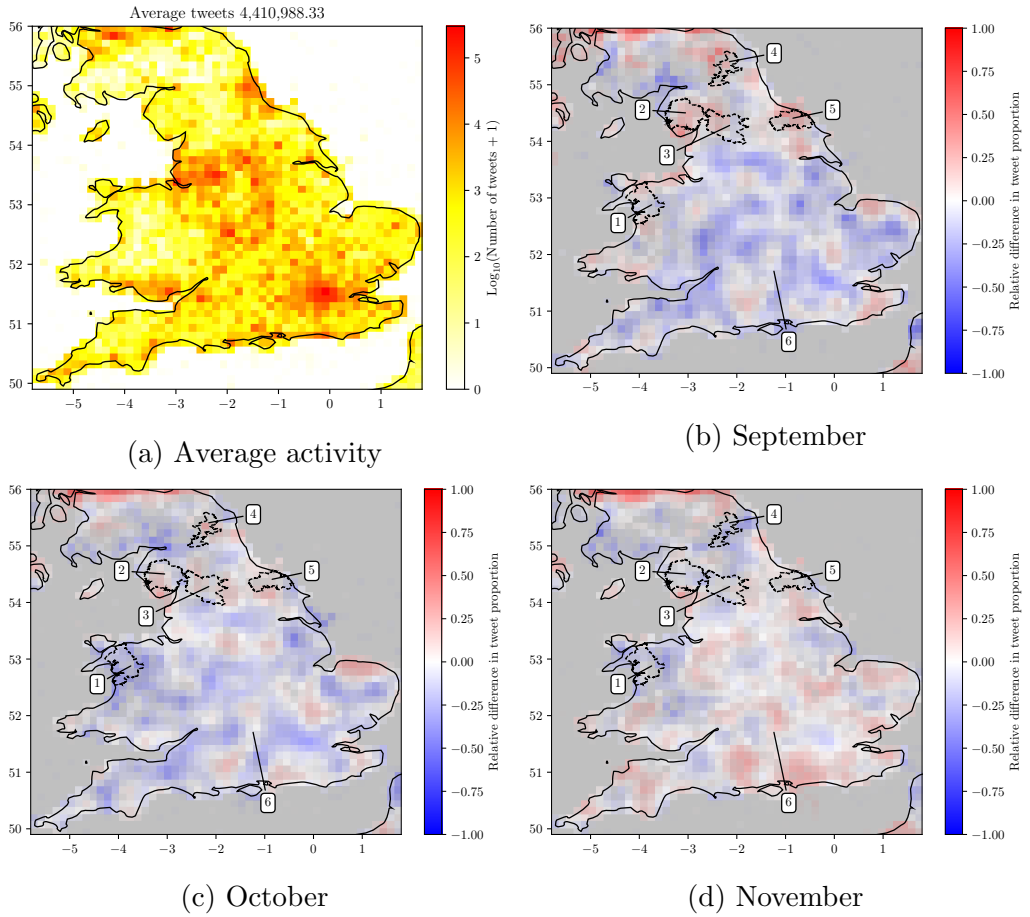
(a) Average activity

(b) April

(c) May

(d) June

Figure 6.1: Smoothed heatmaps of Twitter location activity in 2020 relative to 2019 for April, May and June. Red cells indicate a greater proportion of the total Twitter activity was observed in 2020, whereas blue cells indicate more activity in 2019. Colour intensity indicates the relative difference of the values between 2019 and 2020 and opacity is determined by the average activity between the two years. White cells indicate limited difference between the two years, brighter colours indicate a proportionately greater difference and grey cells indicate limited Twitter activity in the area.



(a) Average activity

(b) September

(c) October

(d) November

Figure 6.2: Smoothed heatmaps of Twitter location activity in 2020 relative to 2019 for September, October and November. Red cells indicate a greater proportion of the total Twitter activity was observed in 2020, whereas blue cells indicate more activity in 2019. Colour intensity indicates the relative difference of the values between 2019 and 2020 and opacity is determined by the average activity between the two years. White cells indicate limited difference between the two years, brighter colours indicate a proportionately greater difference and grey cells indicate limited Twitter activity in the area.

have considered. One possible explanation for this is that November is unique among the months studied by including fewer holiday periods and highly variable weather conditions disrupting plans for individuals to travel away from their usual areas.

During the autumn period, some areas of England were place under different levels of restriction after the introduction of a tier system at the local level. Over the course of October 2020, most of northern England was moved into increasingly more restricted tiers[2]. These different tier levels limited the meeting of groups in indoor and outdoor private spaces, but under all levels allowed small groups to meet outside. Since these restrictions were typically introduced earlier in Wales than in

[2]Details of the first announcement are available at: `https://www.gov.uk/government/news/prime-minister-announces-new-local-covid-alert-levels`

England, Fig. 6.2c reveals a pattern of reduced tagging away from the urban centres in Wales that stands in contrast to the patterns seen in England. In particular, visitation to Snowdonia National Park has responded quickly, while the English national parks lag behind. In England, Fig. 6.2c reveals that the impacts of these additional restrictions are unclear, although visitation to national park areas appears to be decreasing compared to September.

6.3.2 Individual trajectories

In this section we study user mobility through the trajectories expressed by their successive tweet locations. To more accurately assess where users have travelled, we first separate the tweet pairs into those with evidence of travel and those without. Table 6.1 reports the total number of tweet pairs recorded in each of the 12 months studied. Note that the count for September 2019 is comparatively small due to the collection outage for much of the last 10 days of the month. Given the large number of tweet pairs still observed, the general characteristics of the distribution are likely to be unaffected, but it will affect the likelihood of observing inter-tweet times longer than 20 days.

A very small number of tweet pairs in each month capture tweets posted at the same time (as reported in Table 6.2). Such tweet pairs will introduce degenerate behaviour in the subsequent analysis, and are excluded. Examination of these tweet pairs show no clear pattern explaining the behaviour among the 137 users, although it should be noted that a persistent and frequent contributor to this behaviour is an automated account promoting items on the food waste reduction app OLIO[3].

Not all of these tweet pairs convey travel by individuals during the month. In fact, many of these tweet pairs start and end at the same location. At this point it is worth considering what each of these tweet pairs represent. We are studying mobility of individuals based on the subset of locations they have visited that they choose to make publicly available on Twitter. There is no way of knowing whether other locations have been visited in the intervening period or the route taken for

[3] `https://twitter.com/whatsonolio`

	April	May	June	September	October	November
2019	4,539,277	4,974,003	4,768,512	2,959,338	4,891,612	4,884,956
2020	4,100,897	4,513,837	4,049,674	4,030,118	4,317,527	3,768,414

Table 6.1: Total number of consecutive tweet pairs observed in each month period.

	April	May	June	September	October	November
2019	19	22	15	6	31	22
2020	36	40	47	61	38	30

Table 6.2: Total number of consecutive tweet pairs observed with tweets posted at the same time across the twelve months studied.

any journeys. This makes it particularly difficult to study tweet pairs with the same start and end location in more detail. Since our methodology cannot distinguish between individuals who remained stationary and those that did not tweet during their travel away from the start and end location, tweet pairs with a distance of 0 are excluded from our mobility analyses.

We report the number of tweet pairs with the same start and end location in Table 6.3. We see a clear difference between 2019 and 2020 here with the proportion of tweet pairs not representing travel increasing beyond 90% in periods of lockdown. Under the lessened restrictions in September and October 2020, the rate of tweet pairs representing travel increases, but still does not return to the rate seen in 2019. Despite the methodological caveats mentioned above, we see no reason to expect the lockdown periods to change the tweeting behaviour of individuals, unlike their mobility behaviours. Therefore, this change is the first evidence that the lockdown restrictions have reduced the mobility of individuals in England and Wales.

In Fig. 6.3 we see the cumulative distribution function of distances, times and speeds between successive tweets by users in April, May and June of both 2019 and 2020. We first compare the distributions of the distances travelled between

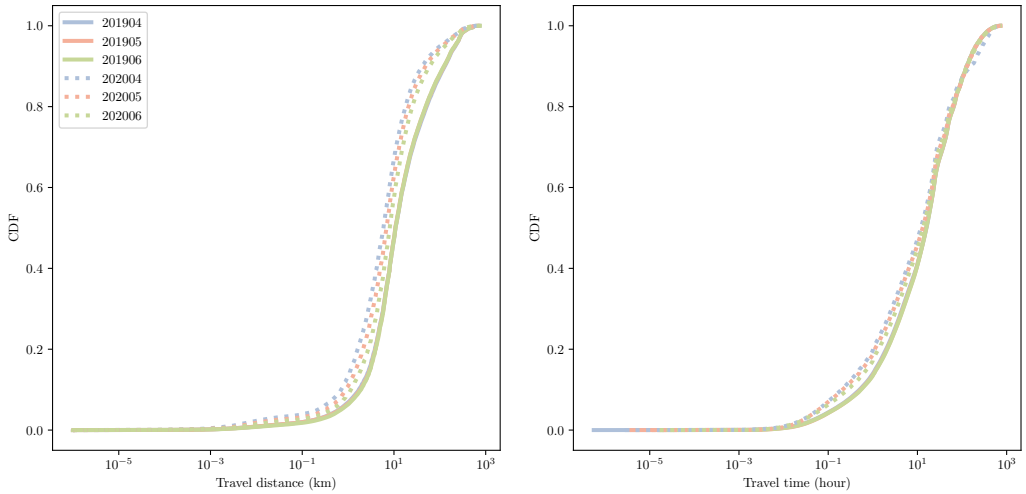
	April	May	June	September	October	November
2019	3,779,743	4,153,814	3,977,985	2,496,989	4,109,807	4,068,772
Prop.	83.3%	83.5%	83.4%	84.4%	84.0%	83.3%
2020	3,890,837	4,247,575	3,756,542	3,576,569	3,870,498	3,460,010
Prop.	94.9%	94.1%	92.8%	88.7%	89.6%	91.8%

Table 6.3: Total number and proportion of consecutive tweet pairs observed with distance 0 between the start and end point across the twelve months studied.

tweets. Fig. 6.3a shows a consistent difference between the distances travelled under normal circumstances in 2019 and under the pandemic restrictions in 2020. The first observation we can make is that there is little difference in the distributions between months in 2019. In light of this, comparing 2019 and 2020 we see that the pandemic restrictions have affected the distances between successive tweets by skewing the distributions towards shorter distances. We see in Fig. 6.3b that the lockdown restrictions also have an impact on the distribution of times between tweets. There is an increase in the likelihood of successive tweets occurring within 10 hours, although after this point the distributions converge. A similar trend but with smaller deviation can be observed in minimum speeds at which individuals must have travelled if the change in their tweet locations reflected an actual journey, as shown in Fig. 6.3c. We see here a small bias in 2020 towards travelling at slower speeds. In calculating the minimum speed required for an individual to have tweeted from the two locations, we notice that a number of journeys record a speed of greater than 1000kmh^{-1}. Such speeds are impossible for most individuals since this threshold is approximately the typical speed for airliners.

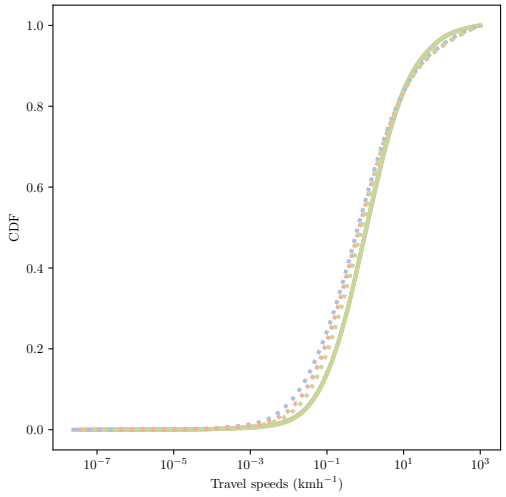
We now repeat this comparison for September, October and November. In Fig. 6.4a we see that the distributions of distances travelled by individuals and the estimates of their speeds more closely resembles pre-pandemic activity but with a very small reduction in distances longer than 10km. The distributions of times and speeds show very little difference between the two years. As with the distribution of tweets across the country in Fig. 6.2, we see more evidence to suggest the (sometimes weaker) restrictions in place in September, October and November have less of an effect than those in place during the spring lockdown.

We compare these distributions in a different way by considering their mean and standard deviations. Table 6.4 reports the monthly mean and standard deviation for distance travelled. Again we see evidence that the periods of full lockdown have reduced mobility as the expected distance travelled is smaller. We also observe a large change in the standard deviation suggesting that travel distances are becoming more consistent. Table 6.5 reports the same statistics for the time distributions.



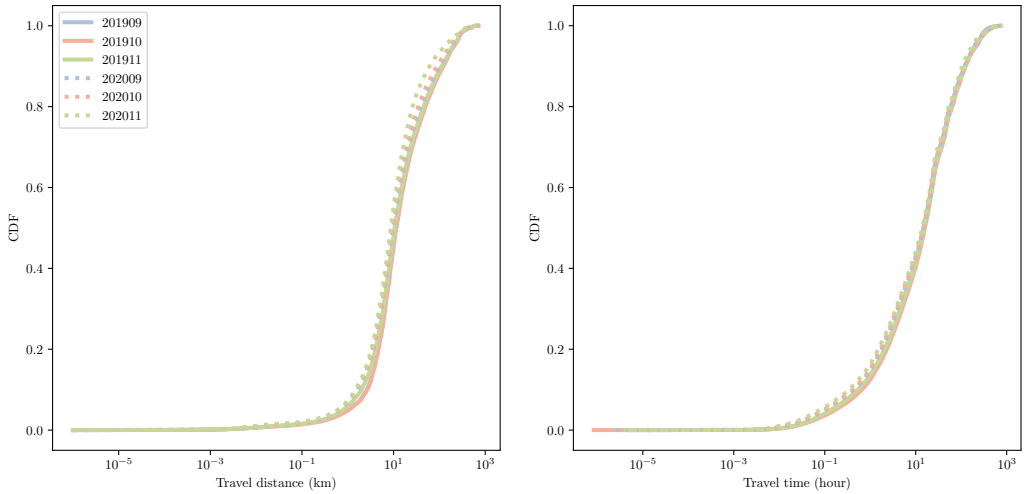
(a) Distances

(b) Times



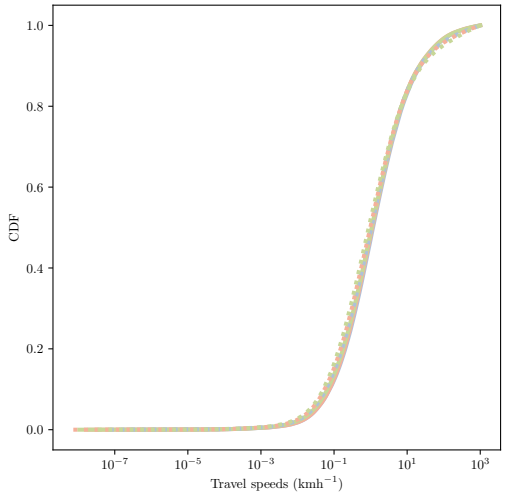
(c) Speeds, truncated at a maximum of 1000kmh^{-1}

Figure 6.3: Cumulative distribution functions for distances, times and speeds for each journey observed in April, May and June of 2019 and 2020. These figures exclude tweet pairs where the distance or time were 0, and where the speed was greater than 1000kmh^{-1}.



(a) Distances

(b) Times



(c) Speeds, truncated at a maximum of 1000kmh^{-1}

Figure 6.4: Cumulative distribution functions for distances, times and speeds for each journey observed in September, October and November of 2019 and 2020. These figures exclude tweet pairs where the distance or time were 0, and where the speed was greater than 1000kmh^{-1}.

	April		May		June		September		October		November	
	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
2019	41.9	78.3	41.0	77.2	40.9	76.9	40.4	75.7	41.2	76.1	38.6	73.6
2020	21.4	55.2	23.4	56.4	26.3	57.8	36.6	72.3	32.9	66.5	27.2	58.9

Table 6.4: Mean and standard deviation of distances travelled (km) in the observed journeys. These values exclude tweet pairs with 0 distance or time and those with speeds greater than 1000kmh^{-1}.

	April		May		June		September		October		November	
	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
2019	45.4	79.3	45.6	80.0	45.5	79.9	42.9	74.5	45.3	79.5	44.5	78.2
2020	49.5	98.6	44.1	81.2	45.0	83.0	43.0	76.0	43.4	78.6	38.9	69.0

Table 6.5: Mean and standard deviation of times (hours) between consecutive tweets. These values exclude tweet pairs with 0 distance or time and those with speeds greater than 1000kmh^{-1}.

Through these metrics there appears to be limited impact from the lockdown events, although it should be noted that the standard deviation of inter-tweet times in April 2020 is unusually high, which may indicate a disruption to existing routines. Finally, we detail the mean and standard deviations of speeds observed in Table 6.6. We see here a universally increased variance in speeds in 2020. A small increase in the mean speed is also observed, but since this is small relative to the increased variance it should not be considered conclusive evidence of increased speeds in 2020.

6.3.3 Anomalous behaviour

In the previous analyses, we have excluded the journeys in which speeds greater than 1000kmh^{-1} are required for the locations to be accurate. In Table 6.7 we see the number of such journeys in each month studied so far. In general, the proportion of these anomalous journeys increases over time, with a particularly high rate in the autumn of 2020. Examining them in more detail we find that many of these implausibly fast journeys occur when the time between tweets is very short. This

	April		May		June		September		October		November	
	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
2019	15.6	67.9	14.6	63.9	14.9	65.5	16.1	69.6	15.2	66.7	15.4	67.3
2020	21.1	88.6	23.1	95.0	22.4	93.5	19.3	85.1	19.9	86.3	22.4	93.7

Table 6.6: Mean and standard deviation of speeds (kmh^{-1}) travelled in the observed journeys. These values exclude tweet pairs with 0 distance or time and those with speeds greater than 1000kmh^{-1}.

	April	May	June	September	October	November
2019	10,642	11,422	12,496	10,335	18,668	21,279
Prop.	0.23%	0.23%	0.26%	0.35%	0.38%	0.44%
2020	10,683	19,811	23,488	30,345	39,783	35,838
Prop.	0.26%	0.44%	0.58%	0.75%	0.92%	0.95%

Table 6.7: Total number and proportion of journeys observed in each month period with a minimum speed of 1000kmh^{-1}. This number excludes any instances of consecutive tweets with the same timestamp.

is not completely surprising since the maximum possible distance for travel in the region of study is approximately 850km, and therefore the time must be less than one hour in order to record a speed of greater than 1000kmh^{-1}. In fact the times are generally much shorter than this as 230,351 (94%) of the 244,790 such journeys capture a period of less than six minutes. We may also ask how the different types of location data affect the likelihood of a journey displaying this improbable behaviour. A total of 25,017 place-tag to place-tag tweet pairs, 217,170 GPS to GPS tweet pairs, 1,722 GPS to place-tag tweet pairs and 881 place-tag to GPS tweet pairs are observed. Since the majority of implausibly fast journeys exist between two GPS locations, we should recall the relative number of such location types in the dataset. As mentioned in Section 6.2.2, only 6.2% of tweets in the dataset include a GPS location and they are therefore disproportionately highly represented among the journeys implying implausibly fast travel. We have no further means of assessing the accuracy of these locations, but note that it is possible to provide inaccurate location data through virtual private network (VPN) services or other third party tools. Since the Twitter API includes more information for tweets including a place tag, we can explore the characteristics of the journeys between two place-tags in more detail. In order to test for inaccuracies in the bounding boxes assigned to location tags in Twitter, we check for any overlap between successive locations. Of the 25,017 place-tag to place-tag journeys with a speed greater than 1000kmh^{-1}, the two bounding boxes intersected in 1,057 cases. In such a situation it is possible that a change to another local place tag has occurred without any offline movement, but since this only accounts for approximately 4.2% of these journeys, it is likely not the defining characteristic of these cases.

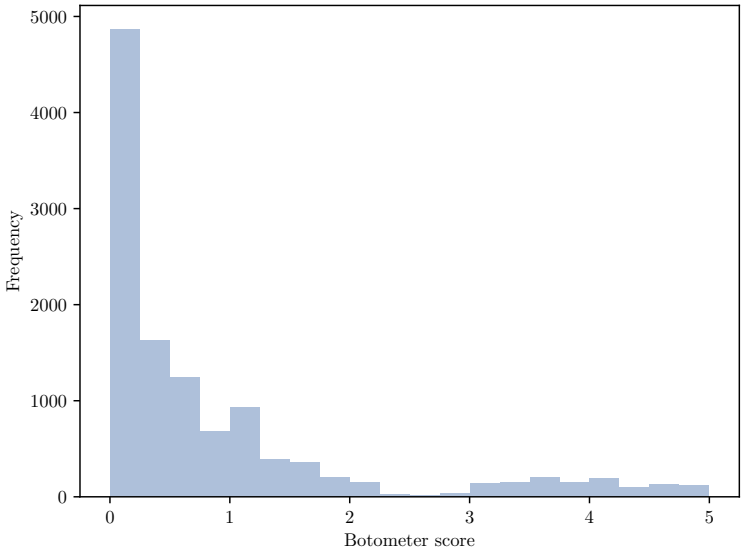


Figure 6.5: Histogram of ratings returned by Botometer for 11,741 of the 13,326 accounts tested. A score of 0 should be interpreted as unlikely to be a bot, whereas 5 indicates likely to be a bot.

One final characteristic of these tweet pairs that can be explored is the possibility that they are arising from inauthentic user behaviour by bot accounts. Pooling cases where two tweets are posted at the same time in different locations with those where successive tweet locations suggest the user was travelling implausibly fast, we flagged 18,622 users with anomalous behaviour. We passed 13,326 of these accounts to Botometer[4] through its API to test the likelihood that they were automated accounts. Of these accounts, 1,585 were unavailable (typically through account deletion or suspension) and the scores for the remaining accounts are shown in Fig. 6.5. This reveals two aspects of the user accounts captured. First, the majority of users tested were found to be unlikely to be bots. Secondly, there exists a slight increase in higher bot scores compared to the 2-3 range. These trends do not suggest that the tweet pairs displaying questionable behaviour are caused by automated accounts and imply that further exploration of the users and tweets in question is required.

[4] `https://botometer.osome.iu.edu/`

6.3.4 Attention networks

To round out the analysis of how Twitter has responded to the events around lockdowns, we compare how the network of Twitter users and the topics they find interesting change over the buildup to and start of two lockdown periods. We compare a two month period between 1st March 2020 and 30th April 2020 and another two month period between 1st October 2020 and 30th November 2020 across three measures to capture the structural changes that occur in the networks from day to day. Fig. 6.6 shows the daily timeseries for bipartite modularity, expected community size and proportion of hashtags that contain our selected terms of interest (*coronavirus*, *covid*, *lockdown*, *pandemic*, *epidemic*). In Fig. 6.6a, we see a decline in modularity and an increase in the proportion of terms of interest over the course of March which demonstrate that the development of the pandemic across Europe is having an impact on the rate with which it is discussed (revealed through the changing frequencies of related terms), and potentially drawing together disparate groups online alongside this common focus (revealed through changes in community size). Two peaks in the importance of coronavirus related terms emerge at the 16th March 2020 and the 23rd March 2020. These two dates coincide with announcements of different levels of measures to combat the spread of the virus, initially in an advisory capacity (16th March) before becoming new government regulations with more severe restrictions (23rd March). We also see that the trends for modularity and the topics of interest over the two months presents a quick change to the event mode before a more gradual return towards the baseline level. The course of expected community size over the two months suggests that it was only the announcement of the more severe lockdown restrictions that led to an increase in the average community size.

If we now look at the trends around the second lockdown in Fig. 6.6b, we again see several days in which the modularity, expected community size and proportion of terms of interest all spike. The changes are focused on two main dates: 31st October 2020 and 7-9th November 2020. As with the previous period, these two dates are tied to major news events and announcements on the course of the pandemic. On the

(a) March - April 2020 (b) October - November 2020

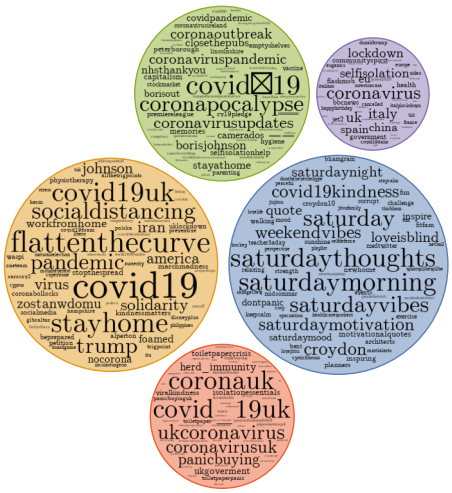
Figure 6.6: Barber's modularity, expected community size on the giant component and proportion of all hashtags in the network containing the terms of interest: *coronavirus*, *covid*, *lockdown*, *pandemic* and *epidemic*.

31st October, UK Prime Minister Boris Johnson announced the period of the next lockdown, beginning on 5th November. The first results of vaccine trials began to emerge on 9th November and signalled a major change in the course of the pandemic to many people. The US presidential election and ongoing results announcements also saw major developments in this period with international relevance. In contrast to the first lockdown, changes in the network are typically smaller and shorter lived for the second lockdown, and return to close to the baseline level within several days.

In addition to the network level statistics, we can use the hashtags present in

the different communities to visualise how important certain topics are on particularly interesting days. Fig. 6.7 presents the hashtags used on six days highlighted as anomalous in Fig. 6.6. In Fig. 6.7a, we see that while the usual trends of weekly Twitter games are still visible in the largest community, each of the other four communities are focused on the growing pandemic. The distinction between these communities is not immediately obvious, but themes like European news (*Italy*), UK news (*covid_19uk*) and general updates (*coronavirusupdates*) do emerge. By 23rd March (Fig. 6.7b), the day of the lockdown announcement, all of the five largest communities are focused on the pandemic but through more distinct lenses such as social distancing (blue), the lockdown announcement (yellow), stay at home instructions (green), criticisms of the situation (purple) and political events or news coverage (red). This pattern of different lenses remains on the 26th March (Fig. 6.7c), with the "Clap for our carers" initiative to show thanks to healthcare staff for their work under extraordinary circumstances.

The prominent terms for the autumn show that the coronavirus topic has become less visible in the largest communities. On 31st October (Fig. 6.7d), the announcement of the second full lockdown only appears in the fifth largest community, where the largest communities are more focused on reality TV (blue), international rugby union matches (yellow), the death of James Bond actor Sean Connery (green) and Halloween (red). Coronavirus related terms are almost completely pushed out of the largest communities on 7th November (Fig. 6.7e) as more results from the US presidential election are released. Despite the dataset requiring tweets to be located in a bounding box around the UK, this internationally relevant event has had a clear impact on the topic structure of Twitter. Finally we can see in Fig. 6.7f that the news about the vaccines is important enough to enter the five largest communities, although only in the smallest, and each of the other four communities are focused on topics away from the pandemic.

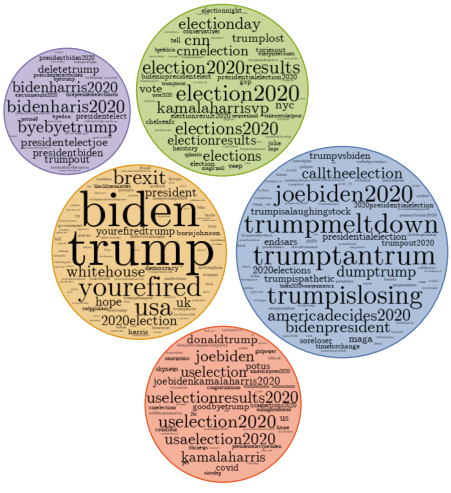


(a) 14th March 2020 (1,273/6,633)

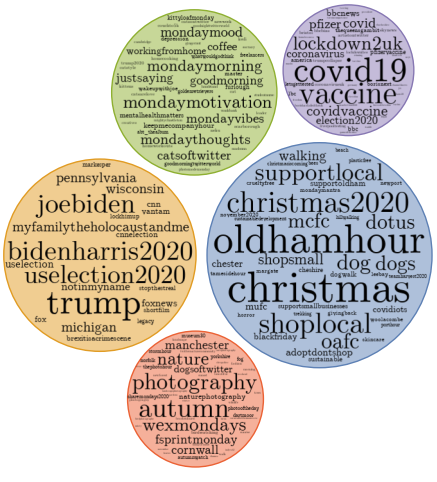
(b) 23rd March 2020 (1,376/6,973)

(c) 26th March 2020 (2,742/8,914)

(d) 31st October 2020 (2,431/7,864)



(e) 7th November 2020 (1,529/7,391)



(f) 9th November 2020 (958/4,798)

Figure 6.7: Hashtags and proportion of nodes appearing in the five largest communities on selected days. Circle size is proportional to the total number of uses of hashtags in each community and numbers in the captions refer to the number of nodes in the five largest communities/number of nodes in the giant component.

6.4 Discussion

In the previous section we have seen a number of ways in which the patterns of behaviour among Twitter users from England and Wales have responded to the different levels of restriction in place to combat the spread of the COVID-19 pandemic over the course of 2020.

We compared the spatial distribution of Twitter activity under typical conditions in 2019 with different levels of lockdown restrictions in Section 6.3.1. Comparing the same months across the two years we see that the "stay at home" advice and wider social distancing measures led to a systematic reduction of activity across much of the country, and particularly in those areas that would normally see visits for recreational activities that were no longer permitted. This pattern evolved across April, May and June during which time the restrictions were most severe, with only a small number of circumstances in which individuals were permitted to leave their homes. In the autumn of 2020, restrictions were at their weakest, and typically different at the local level, following a summer of relatively few cases. During this period, Fig. 6.2 shows that the patterns of lower activity in 2020 in sparsely populated national park destinations has been reversed. October (Fig. 6.2c) also highlights some of the impact of the local restrictions and in particular the "circuit breaker" temporary lockdown introduced in Wales from 23rd October as areas of the north east and Wales display the behaviour more commonly seen in periods of lockdown.

Through these comparisons we have shown that the measures in place to combat the spread of coronavirus across England and Wales have affected the locations users choose to include in their tweets. The patterns appear to be consistent with fewer posts from leisure locations. This trend is reversed in periods of lessened restrictions, potentially indicating that the desire for comparatively safer holidays within the countries saw many people take advantage of the relatively relaxed restrictions in September and October. As a result we can say with some confidence that in answer to RQ1, attention to different locations is affected by periods of lockdown, but the

strength of the effect is determined by the typical reason for tweeting from the location.

To expand on the difference in location distributions under the lockdown conditions we measured the mobility of users through their locations in successive tweets. By measuring the distance and time between these tweets locations, we can begin to understand the dynamics of the Twitter community across England and Wales during the study period. Many of these tweet pairs involved no movement, although the various restrictions increased the proportion (Table 6.3). In Fig. 6.3 we saw that the lockdown restrictions in place for April, May and June 2020 led to a higher likelihood for the distance travelled and time between tweets to be shorter than the same months in 2019. As with the location distribution, Fig. 6.4 showed that the relaxed restrictions in the autumn of 2020 had a lesser impact on the mobility of Twitter users. Only the distributions of distances between tweets were notably different in 2020 compared to 2019, and provide further evidence that the more severe restrictions affect mobility to a greater extent.

In response to RQ2, the user trajectories described by successive tweets show that there are some changes under periods of lockdown, although it appears that the size of the change is dependent on the severity of the lockdown restrictions. One other possible explanation for the difference in responses between the two lockdown periods is a level of acclimation effect as individuals adjust their travel habits to produce similar mobility patterns on Twitter, but in alternative COVID-safe ways.

The final component of our analysis was an investigation of how the structure of the network of discussion topics on Twitter changed during the anticipation and introduction of the two full lockdowns in the study period. In Fig. 6.6 we measured this change through Barber's modularity, the expected community size and the proportion of hashtags that included a series of terms central to the coronavirus pandemic. We saw that following the various news events of interest to the general public, the network structure responded by becoming less modular, with an increase in the expected community size, and where these events were pandemic related an increase in the proportion of hashtags containing the terms of interest. Here

again the patterns are different between the spring and autumn lockdowns. In the spring lockdown, the announcement of the escalating severity of restrictions leading up to the lockdown starting on 26th March causes a comparatively large and persistent change in the modularity and visibility of pandemic terms in the network. By the autumn lockdown, changes in modularity and hashtag visibility become comparatively smaller and last only days before returning to the baseline. While the characteristics measured here cannot be as clearly linked to the levels of social restrictions at the time, we can see that the social media ecosystem in England and Wales reacts differently to these major events.

To better understand the focus on the notable days highlighted by the network-level trends in Fig. 6.6, we examined the hashtag composition of the five largest communities on a series of important days. Fig. 6.7 again showed that reactions to the pandemic changed between the spring and the autumn. In the spring, most of the five largest communities are organised around terms covering different aspects of the pandemic. By the autumn however, this visibility has greatly diminished as other topics attract large and coherent conversations. With reference to RQ3, we find similar event trends during the first lockdown as those observed by Borge-Holthoefer et al. [26] for the 15M movement. This trend was not maintained into the second lockdown period however, suggesting again that some adaptation to the event has taken place.

Through each of the three different facets of life in England and Wales as reflected in Twitter we have seen that the reactions to pandemic-related events in the autumn were different to similar events in the spring. This difference may be a manifestation of the ongoing pandemic becoming incorporated into daily life and therefore no longer exceptional, or even "pandemic fatigue" [14], where the ongoing situation impacts on the mental health of individuals. The transition of the pandemic from exceptional occurrence into everyday life explains the trends seen in Section 6.3.4 as the regularity of pandemic-related events may limit their ability to generate sufficient collective attention to form large and cohesive communities compared to other, more novel, topics. The manifestation of different topics online and the mobility inherent

therein are likely driven by different processes however. Similarities between the mobility patterns in autumn 2019 and autumn 2020 as seen in Fig. 6.4 imply that the restrictions in place on offline mobility have very little impact on the mobility observed via Twitter. Considered in the context that the restrictions in place for much of November 2020 were broadly the same as those for April, May and June, this suggests that the restrictions in November may have been less effective than those in the spring. With Twitter data alone we have no way of knowing if this is a result of reduced compliance with the social distancing restrictions, but the changes in which activities were permitted may be enough to manifest the similarities seen. The key relaxations that are likely to influence the similarity of the mobility distributions in November are schools remaining open and dispensation for exercise within the local area. Since the Twitter user base tends to be younger than society as a whole [3], schools remaining open regardless of other restrictions replaces a degree of normal mobility that was missing during the spring lockdown. Increased exercise outside the home may have a similar (although likely smaller) effect. These additional journeys are likely to be shorter however as restrictions still asked individuals to stay local to their home. Since much of the difference in distance distributions seen in the spring lockdown was for shorter journeys (Fig. 6.3a) local mobility is likely the key differentiating factor.

Beyond this study of how attention to the pandemic has affected Twitter, the mobility analysis in Section 6.3.2 raises questions about the accuracy of locations included in tweets. As seen in Table 6.7, a small but persistent proportion of these tweet pairs require travel at impossibly fast speeds to be accurate. Of more concern is how the overwhelming majority of these journeys are between two tweets located by GPS coordinates rather than a bounding box. These types of location are typically considered as more accurate due to their higher precision, but we have seen over 200,000 tweet pairs that suggest at least one of the locations are inaccurate. It may be possible for these locations to be provided by shared accounts where individuals have tweeted from different locations in quick succession, but given the quantity of tweet pairs, it seems likely that some misleading locations have been provided

through the various means of location spoofing that exist.

Looking to the future, this work poses two additional questions. We've seen how patterns in England and Wales have been affected by the different levels of restrictions, and this work would be enhanced by studies in other countries with different restrictions in place at different times. Such additional comparisons will enrich the relationship between online communication and offline restrictions seen here by potentially highlighting which types of restriction produce which effects and whether the transition of the pandemic from exceptional to routine extends to other cultures. Particularly valuable insights could be drawn from similar analysis of countries with no official restrictions to get a sense of how management efforts decided by individuals manifest at the country level.

The second question concerns additional analysis of the implausibly fast journeys highlighted in Section 6.3.2. We've seen in Fig. 6.5 that it is unlikely that bot accounts are driving this behaviour and understanding the users and the circumstances of their posts will go a long way towards addressing the potential concern in the accuracy of locations in tweets. Verifying the accuracy of geo-tagged tweets will need external data and cautious analysis to respect the privacy of individuals and the Twitter terms of service. Without collecting additional data, analysis of the tweets including location tags is possible through understanding the relevance of the location given to the text in the tweet. Matching names or those of nearby locations is one such method, but more sophisticated techniques could make use of framing in the text to indicate location accuracy. A successful implementation of this approach would be able to distinguish "I wish I could have gone on holiday to London", "I had a great time in London today" and "Throwback to the best holiday to London last year" and rate the location accuracy of the second example as the highest. Another aspect of the location behaviour that has not been considered here is the source platform and whether the split between mobile, desktop or third-party services biases the likelihood of observing these locations with questionable accuracy.

Similar attention on different scales?

What does the previous chapter tell us about collective attention?

In the previous chapter, we used the COVID-19 pandemic and associated extended lockdowns as a case study for a longer-term disruptive influence on collective attention. Using an expansive dataset of location-tagged tweets from across England and Wales, we investigated how locally relevant events in the ongoing coronavirus pandemic shaped the online conversation on Twitter.

By comparing the spatial distribution over time, we saw that the lockdown restrictions had varying impacts on the tagging of different locations in tweets. During the first lockdown, we saw reductions across much of the country, with particular declines emerging in sparsely populated areas primarily used for leisure purposes. In contrast, the autumn lockdown showed much smaller year on year variation in location tagging, suggesting that some acclimation had occurred or other means of enjoying these areas had emerged.

Chapter 6 also leveraged the locations users included in their tweets to measure changes in mobility at different stages of the pandemic. As we saw with attention to different locations, we observed a greater change in mobility from the previous year in the first lockdown compared to the second. Our mobility analysis also revealed a number of trajectories indicating the users travelled implausibly fast. This surprising finding raises the question of how accurate these locations actually are. Analysing

these users in more detail found explanations under some circumstances, but more work is required to fully understand how this behaviour has emerged.

Returning to the study of community structure, this time in reference to user-hashtag networks, we found that content patterns do show some changes. These changes peaked early in March 2020 after a critical awareness of the pandemic had been reached following the rapidly growing case numbers in Italy at the time. We showed that during the first lockdown, notable news stories led to coronavirus terms becoming common in many of the largest communities and a series of structural changes that were ultimately short-lived. Once again, we observed diminished attention responses to the onset of the second lockdown, reinforcing the observation of an acclimation effect making news events in the ongoing pandemic less worthy of attention.

Our work studying the coronavirus pandemic adds some new perspectives on how collective attention varies over time. In contrast to the work in Chapter 5, we are able to consider two new facets of attention that are typically difficult to capture. Firstly, we are able to see how collective attention changes with extended events. Secondly, we are able to understand how collective attention responses to repeated, similar events differ. We have certainly seen diminished responses as the pandemic continues, but future work should consider whether similar trends exist in other topics, and over which timescales.

How does this inform our upcoming work?

As noted before, the previous two chapters have considered collective attention responses to events occurring over varying timescales, either weeks in the case of the Paris Agreement announcement, or months in the case of the COVID-19 pandemic. In order to more directly compare these two events, we need some way to understand how collective attention is developing over time.

Attention as a process has seen many attempts to model its effects on network formation (e.g. [18, 75]), but these models don't enable direct comparisons to be easily made across different events. Given the scale-free nature of most social net-

works, and the size of social media datasets, it is possible that collective attention events can capture anywhere from hundreds of users to hundreds of thousands of users. We can also see events lasting from minutes to days and occurring at any time of the natural daily (or weekly) rhythms of social media. Across such situations, traditional timeseries methods will be unable to resolve the differences to reveal whether similar generating processes are influencing the development of these social networks.

In the next chapter we aim to address this challenge by introducing a standardisation process that allows for comparisons across time and volume scales and taken at different times of natural rhythm of social networks. We do this through two steps. First we change our consideration of time to reflect the messages we observe in our dataset rather than real-time, which compresses quiet periods of time and stretches busy periods. The second step tracks the development of the topic of interest as a proportion of the observed messages so far. We test this new method on a series of hashtags and their likely events drawn from a long-term Twitter dataset of the ongoing Brexit process in the United Kingdom, and attempt to find similar characteristic behaviours to those outlined by Lehman et al. [111].

Chapter 7

Shapes of collective attention online

Abstract

Online social media continues to be a cornerstone of communication, and news media consumption. Unlike traditional media, the magnitude and timescale of attention different social media events generate is wildly divergent, allowing viral content to reach millions of users over the course of hours or weeks while other content is ignored. Understanding the characteristic "shape" of a peak of attention on social media—that is its duration, size and dynamics—can point towards different generating processes and predict which items will become popular. Studies of collective attention have typically imposed a timescale on the data through binning that may hide nuances in the underlying processes. Here we address this limitation by presenting a standardisation process that does not mandate a characteristic timescale, allowing for direct comparison of events at vastly different temporal or activity volume scales. This standardised representation highlights several characteristic profiles which we argue signify different types of attention patterns and underlying user behaviour, universal across timescales and magnitude. We find that these shapes of attention form a continuum with clear structure but no discrete categories, but certain parts of the continuum can be used to infer particular mechanisms of attention flow.

7.1 Introduction

Following an explosion of popularity over the last two decades, online social media services have become important components of social networking and media dissemination. Measuring the activity of individual users is relatively straightforward due to the wealth of data available through social media platforms, but collective attention to a particular topic, event or piece of media remains more difficult to observe. Collective attention has grown as a topic of interest, though there is no consensus on a formal definition. Wu and Huberman [199] consider collective attention to be the process of how attention to novel topics propagates and eventually fades among large populations. This definition neglects to consider the source of the attention and whether it is truly collective, that is to say driven by interactions between individuals. Under Wu and Huberman's definition, a large group of individuals simultaneously, but independently, paying attention would be considered collective attention, despite the "collective" component influencing nothing more than the number of members in the system. Lin et al. [114] propose shared attention as "a temporary state in which the individual members of an audience for an event are mutually aware of each other's attention to the event", a definition which incorporates such a requirement. Given the difficulty in measuring this second definition, it is likely that the debate over a formal definition will continue. For the purposes of this manuscript, we impose our own definition of collective attention that draws themes from both of these existing definitions. We define collective attention as the state where the attention of others increases an individual's attention to a topic. In this way, we are incorporating a measure of preferential attachment among attention as discussed in the context of network formation by Barabási and Albert [18], and explicitly require a level of interaction between participants in the system.

Here we will outline and test a new method of understanding collective attention in social systems through the use of a large dataset of conversations around Brexit, gathered from the social networking site Twitter. This paper is organised as follows.

Section 7.2 will outline previous work in this area including previous attempts to characterise collective attention and identify some of their potential shortcomings. In Section 7.3 we give more details about the Brexit dataset and present full details of our new methodology designed to address these shortcomings. In Section 7.4 we demonstrate that our methodology produces meaningful representations of different activity patterns and outline the characteristic shapes we discover. In Section 7.5 we discuss the implications of our findings, before outlining the next steps for this method in Section 7.6.

7.2 Background

7.2.1 Collective attention in online platforms

Timeseries methods have long been a popular tool to study periods of increased attention (e.g. the 2012 US presidential election debates [115] and reactions to online news stories [37]). Previous work to measure collective activity has examined temporal patterns and decay rates of interactions on the news aggregator website Digg [199]. By measuring the rate of decay in Digg interactions with a number of stories over time, Wu and Huberman were able to demonstrate it followed a pattern of stretched exponential relaxation with a half-life of approximately one hour. Other approaches apply user metadata to predict adoption of hashtags (user-defined tokens signifying particular topical relevance) [202] or content metadata to identify potentially malicious attempts to spam the online conversation [109]. Many studies use measures of divergence from some baseline to understand systemic change in social networks with some success at detecting and quantifying collective attention events [161] and capturing the evolving phases of attention under some external stimulus [177]. In addition to considering only macro-level variations, Lin et al. [114] and De Domenico and Altmann [49] show that different types of activity such as retweeting, replying and sharing on the social media platform Twitter respond differently to increased collective attention.

Other attempts to measure collective behaviour make use of the meso- and

macro-scale tools of network science. Network structure influences the content users are exposed to, with recommender systems which suggest connecting with users or content based on a user's immediate network neighborhood. For example, by recording the followership of users, it is possible to enhance the prediction of popular memes [194]. He and Lin [84] study the social media platform Twitter by building networks of successive hashtag usage and find that the topology of the underlying network changes around disaster events. Another common approach considers the connectivity between users and other objects in a bipartite network formulation to track how users engage with content over time [35] and how topical clusters evolve (e.g. [158], [118]). Such studies provide insights into how different types of content respond to increased attention, for example demonstrating that negative aspects such as controversy [74] can become more prevalent in such circumstances. Network methods have also been shown to be useful for event detection purposes [132].

Beyond the techniques used to measure collective attention online, it is common for studies to propose different modelling approaches to investigate the factors that influence the growth of attention at scale. Given the similarities between information diffusion and epidemic spreading, adaptations of the well-known SIR (susceptible-infected-recovered) and SIS (susceptible-infected-susceptible) models for collective attention [100] have been applied in addition to models inspired by cellular growth [183]. Agent-based models are frequently applied to test the importance of user-level effects such as neighbour behaviour [75] or memory of recently seen content [22]. Some models consider network growth in discussions (e.g. [123]) through extensions of the concept of preferential attachment proposed by Barabási and Albert [18]. Others still use a variety of mathematical modelling frameworks to capture how information and attention spread, such as: Markov chains [9, 67]; differential equation models of information flow [81]; competition between topics for the same limited attention [192]; and size and lifetime of retweet cascades [31]. It has been shown that such models do not need to be that complex to mimic observed behaviour at the system-level. Huynh et al. [90] developed a model using only two parameters: virality of the content, and speed of decline after the peak, and are still

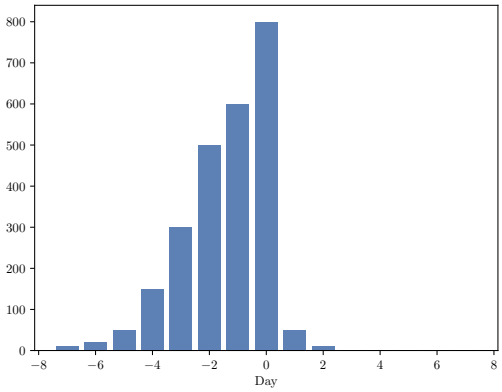
able to capture the real-world phenomena.

Understanding the processes underlying collective attention presents some valuable opportunities to researchers for understanding the rising trends of disinformation in modern society. Vosoughi et al. [182] compare the cascade statistics of real and fake news stories, ominously finding that fake news stories spread further, faster and to more unique users than real news stories. Zhao et al. [209] complement this by demonstrating that the spreading topologies of real and fake news are characteristically different. On the other hand, Mitra et al. [130] exploit these differences to show that veracity classes can be identified by information activity over time. Social networks provide fertile ground for organising social action and protests and have already been studied around the Arab Spring [173], political rallies [57] and climate change protests [166]. Such advances will be invaluable for tackling issues arising in a connected society, but ethical consideration needs to be given to ensure that these tools do not adversely affect the rights of social media users.

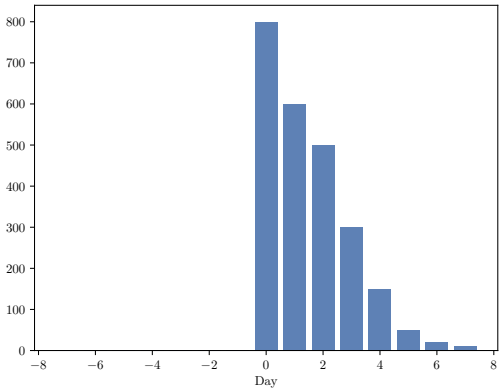
7.2.2 Types of collective attention event

Beyond these attempts to measure and model attention, some effort has been devoted to classifying different attention patterns. The work of Lehmann et al. [111] is often cited when considering the temporal trends of online activity. They studied the distribution of activity around the day of peak attention and found that events fell into four classes that capture different types of event (see Fig. 7.1 for illustrations of the patterns typical of each class). The first class captures events with a gradual build up of attention until the peak before quickly declining. The second class reverses this trend, with a sharp increase in activity at the peak and a gradual decline afterwards. The third class incorporates both a gradual increase and a gradual decrease in attention around the peak and the fourth class covers events in which the activity is concentrated almost entirely on the peak day. Extensions to this work have been published over the years, such as [90] which supports the existence of these distinct categories.

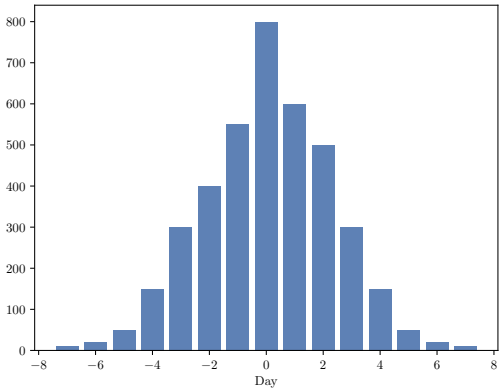
These categorisations are useful for understanding system dynamics but make



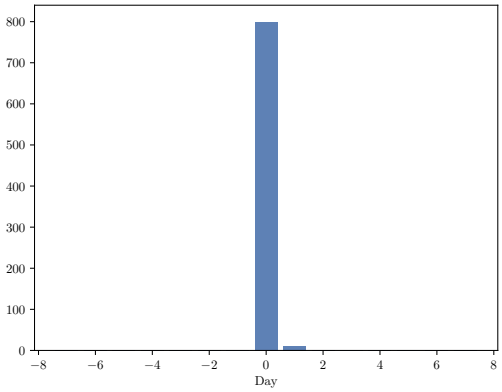
(a) "Before the peak" class



(b) "After the peak" class



(c) "Before and after the peak" class



(d) "Peak only" class

Figure 7.1: Illustrations of the activity profiles represented by each of the four classes found by Lehmann et al. [111]. Each illustration is centered on day 0, the day of peak activity.

a potentially problematic assumption by assigning a characteristic timescale to the process. Lehmann et al. bin observed tweet counts into one-day time slices to form a timeseries, meaning that events occurring within a single day are represented by a single value, regardless of their underlying behaviour. This may be particularly problematic as observation of social media suggests that different topics can be the focus of attention for anything from minutes (e.g. press releases and sporting events) to weeks (e.g. natural disasters and elections). This variation means that the choice of bin size needs to be carefully considered; daily or hourly bins may be the most obvious choices but should be analysed with the caveat that certain features may be over- or under-exposed at the chosen resolution. Furthermore, on certain timescales the day-night or weekly cycles, important patterns in social media activity, are obscured and may require special attention. Avoiding these problems requires manual interventions that are only feasible for investigations of small numbers of events, or those where other information about the events is available.

As a more simplistic measure of collective attention, increased volume of activity can be seen as increased attention to a topic. What is not clear, however, is whether increased activity is a sign of many instances of individual attention growing independently under exogenous factors or collective attention driven by the interactions between users. Such a distinction is often impossible to measure using the data available to scholars of online social systems and as a result peaks of activity are often used as proxies for collective attention. Hashtags have frequently been used to identify such activity peaks since they provide a convenient indicator that the user considers their post relevant to the topic. Online social media sees dramatic variability in the total magnitude of activity a certain topic receives which further complicates the relationship between activity and attention. From a mechanistic perspective, there is no way to recognise whether the observed activity volume arises from different sizes of user base, increased baseline activity across topics or changes in interest to each respective topic. There are several avenues to mitigate this effect, such as the work of Eyre et al. [60] who use posting patterns on business Facebook pages to detect downtime after natural disasters in different areas. Their

methodology aggregates data from multiple businesses in the area over time and in order to account for the variation in normal usage patterns of different businesses they use a "probability integral transform (PIT)" technique. The PIT uses the cumulative distribution function (CDF) of daily post rates over a year to estimate the probability that a business makes fewer posts in a day than expected. As a result, the PIT transforms the very different scales of business posts to the interval [0,1] for each business, presenting the secondary benefit of making them more directly comparable and suitable for subsequent aggregation and analysis.

7.2.3 A scale-independent comparison for events

After reviewing the previous attempts to study collective attention in the literature, we identify two problems that need to be resolved, and use them to define the goals for our proposed methodology:

- Timeseries binning - requiring a bin size can over- or under-expose certain aspects of event behaviour depending on their inherent timescales.
- Event comparisons - directly comparing distinct events is hampered by different volume scales and unknown factors affecting baseline activity rates.

These two limitations make it difficult to work with large datasets of collective attention events. While it is easy to bin timeseries at the same resolution, it is difficult to know whether the resolution is appropriate for all events considered. Facilitating direct comparisons on the other hand requires significant effort to normalise both the time and volume scales, as well as address any independent factors affecting activity volumes.

In this manuscript we present a novel methodology for analysing peaks of activity over time, even under the wide range of time and volume scales observed. This new technique will address the two main shortcomings of existing timeseries approaches and presents a valuable capability for comparing dissimilar events to reveal similar generating mechanisms manifesting at different scales. Primarily our methodological improvement removes the need to determine a suitable timescale a priori by

measuring the rate of arrival of a portion of the event tweets relative to the length of the interval. Considering the tweet rate can be adapted to apply a normalisation which also addresses our secondary goal of facilitating direct comparisons between events of different time and volume scales that were previously only possible through representations customised for each event interval. Using this new methodology we analyse a dataset of hashtags used in the Twitter conversation around Brexit, the common term for the United Kingdom leaving the European Union. We demonstrate the effectiveness of our methodology for comparing events at different scales and find that three of the four classes characterised by Lehmann et al. are recovered, but that their "peak only" class is likely an artefact of overly coarse binning of the timeseries. Instead we find a fourth class where activity patterns abruptly shift between two modes, revealing a new behaviour in hashtag usage that is not well-represented in existing methods.

7.3 Methods

7.3.1 Data

The experiments in this paper make use of a dataset collected from the Twitter Streaming API using the list of keywords in Table 7.1, selected to capture a range of opinions and themes about the Brexit negotiations between the United Kingdom and the European Union. The collection runs between 1st January 2018 and 26th June 2019 and includes 12,545,942 tweets. This period included a few short outages in the collection but since the longest covered approximately two days this is not expected to affect the subsequent results. In order to extract signals of attention, we identify the hashtags included in the tweets. In the absence of any specific measures of collective attention, we choose these as specific and intentional indicators of the content and terms that users are not only aware of, but also invested in enough to tag in their own posts. Over the 18 month period we collect 950,508 unique hashtags from 1,333,403 users.

brexit	leaveeu	no2eu	voteleave
euref	betteroffout	voteno	ukineu
remainineu	stayineu	yes2eu	incampaign
referendum	european union	strongerin	eureferendum

Table 7.1: The keywords supplied to the Twitter Streaming API to gather the dataset, returning any use in the tweet text (including hashtags).

7.3.2 Peak identification and dividing the timeseries

For our analysis, we select a sample of 100 hashtags from the 4,278 hashtags used at least 500 times across the 18-month study period. This threshold is low, averaging approximately one tweet per day, but necessary to ensure a minimum level of activity for an event. To consider different events for the same hashtag it was necessary to manually divide the activity for these hashtags into a series of intervals containing either an event or a period of persistent activity. We explored ways to do this algorithmically, but find there are issues with these methods. The natural choice for this activity would be peak detection methods (e.g. [111, 143]; see [85] for further comparisons) or tipping point analysis [28], but there is no guarantee for such methods to identify the complete lifetime of the event (that is, not just the event peak but also its onset and dissipation). We also note that there is no best choice for peak detection algorithms and comparisons of methods have found large variation in accuracy metrics across different topics [85]. Given these limitations, the sample of 100 hashtags were chosen such that event periods of twice the surrounding baseline rate were clearly identifiable in the day binned timeseries of tweet counts through visual inspection. Although this process is not feasible for analysis of large datasets, and biases are introduced by choice of bin, this approach is necessary to demonstrate the efficacy of our proposed method and does recover events over different timescales (as evidenced in Fig. 7.3). See Fig. 7.2 for an example of a timeseries divided into single-event intervals. In this figure, and throughout the manuscript, we number hashtag intervals sequentially across the study period and also record the number of tweets containing the hashtag included in the interval.

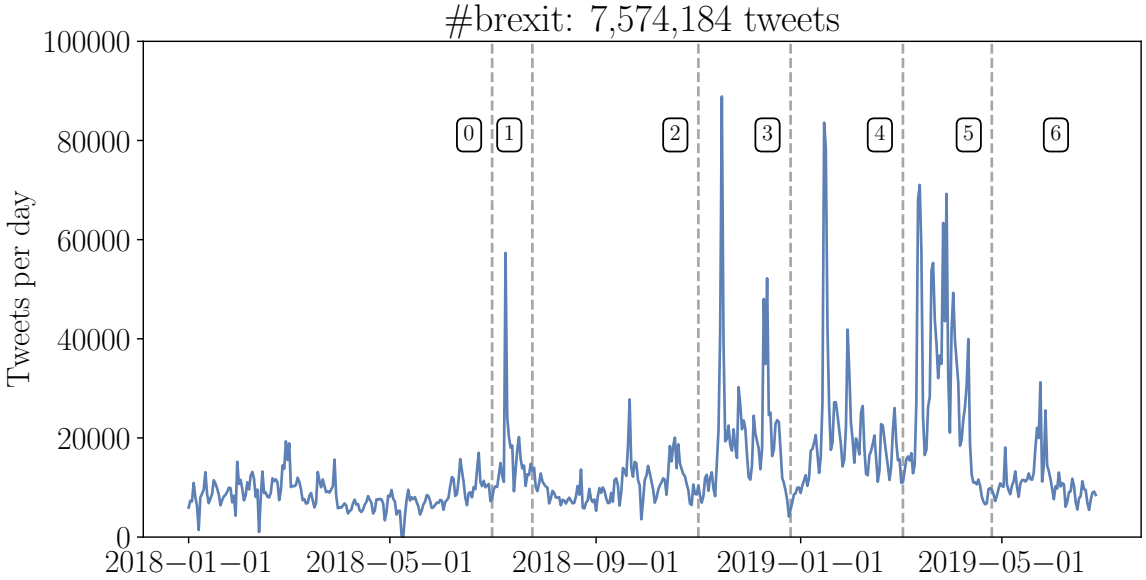
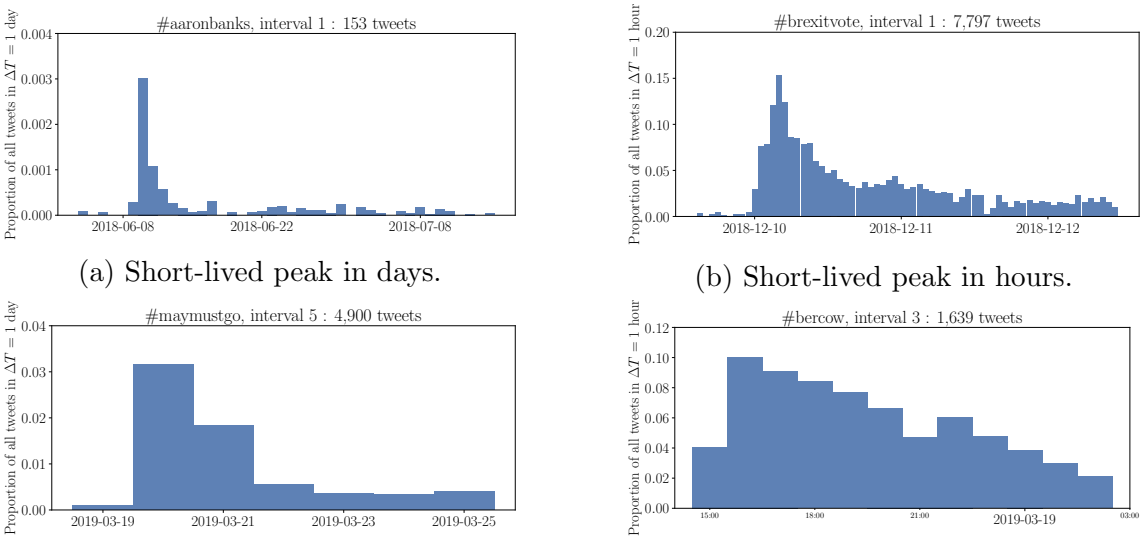


Figure 7.2: Each hashtag is divided into a series of intervals over the study period based on periods of increased activity. In this example, seven periods are defined. The numeric labels for each period are the interval numbers as referenced throughout this manuscript.

7.3.3 Scale-independent representations of timeseries

We seek to transform the activity in each event interval such that we can compare events with different time and volume scales. This is of great importance given the long-tailed distributions of many metrics related to online social networks [17], which imply that events with the same underlying generative mechanism can manifest at dramatically different scales. Traditionally, analysis of events has binned the data at some prescribed temporal resolution, such as hours or days. The choice of bin size can critically alter the visibility of certain trends, and without a consistent timescale for events on social media there remains no obvious choice. Fig. 7.3 demonstrates the similarities between events with different scales that can be observed using examples from our dataset. In light of this observation, we define a fair comparison between two intervals as one that reflects common trends of growth and decay, whether a meme spreads through an isolated group over a few hours or across the globe over many days. To achieve this goal, our methodology will not consider time or absolute tweet counts and instead consider the rate of tweets in a given time interval to avoid the need to bin timeseries at a particular resolution. Considering tweet rate also helps to resolve independent non-event drivers of tweet volume, such as the day-



(a) Short-lived peak in days.

(b) Short-lived peak in hours.

(c) Gradual decline over days.

(d) Gradual decline over hours.

Figure 7.3: Different hashtags can show similar trends when viewed at different temporal scales (defined by bin width). To compare these directly with each other careful normalisation across volume also needs to be applied.

night cycle, week-weekend cycle, and longer-term overarching trends in platform usage. At a suitable scale, binning resolves some of these factors but the difficulty with determining such a scale for multiple events motivates our primary goal. Our approach to manage these factors is detailed in this section.

Before calculating any representations, each tweet is assigned a sequential ID. This enables consideration of rates relative to the collected data which are obscured when using tweet IDs assigned by Twitter (since they are affected by activity globally). Then, given an interval with n tweets over an arbitrary temporal width, the scale-independent representation is derived as follows.

- Collect the dataset IDs assigned to each tweet in the interval that uses the hashtag of interest.

$$(11207549, 11214100, 11215455, ..., 12498903, 12519628, 12523526) \quad (7.1)$$

- Normalise these numeric IDs such that they span [0,1]. At this point, the number for the ith hashtag use indicates the proportion of all observed tweets

in the interval required to see i uses of the hashtag.

$$(0, 0.005, 0.006, ..., 0.981, 0.997, 1) \tag{7.2}$$

- Choose the dimensionality N of the final representation.
- Sample the cumulative distribution function (CDF) of the normalised tweet IDs at $N + 1$ equally-spaced quantiles (including 0 and 1) determined by the chosen dimensionality.

$$(0, 0.042, 0.069, ..., 0.841, 0.939, 1) \tag{7.3}$$

- Calculate the difference between the sampled quantiles, i.e. the proportion of observed tweets required to reach the next quantile.

$$(0.042, 0.027, 0.02, ..., 0.045, 0.098, 0.061) \tag{7.4}$$

- Calculate the tweet rate between quantiles by inverting the proportion of assigned tweet IDs required to reach the next quantile.

$$(23.7, 36.8, 48.5, ..., 22.3, 10.2, 16.4) \tag{7.5}$$

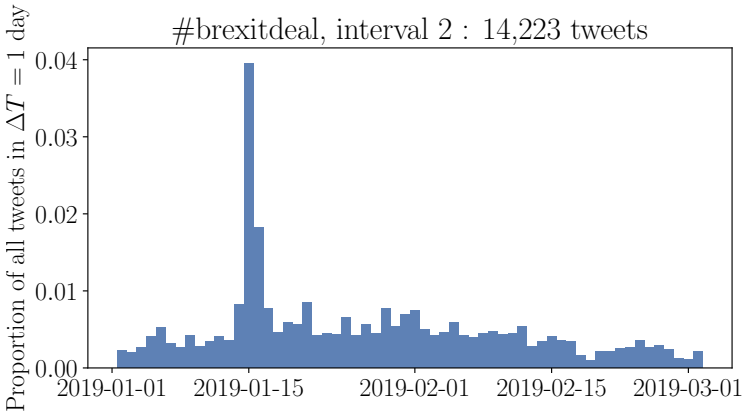
We are left with an N dimensional representation of how the usage rate for the hashtag changes across the interval. More concretely, given the choice $N = 50$, the vector elements in $[0, 1]$ give the proportion of all tweets observed in the time period required to give the next 2% of total hashtag uses, i.e. the rate of all tweets that include the hashtag we are investigating. We invert the values in the representation such that larger values indicate a higher frequency of tweets containing the hashtag of interest to aid interpretation.

To demonstrate this approach in practice, Fig. 7.4a shows the fraction of all tweets each hour containing the hashtag *#brexitdeal*. We transform the timeseries into cumulative usage of each hashtag against all tweets in the dataset, shown in

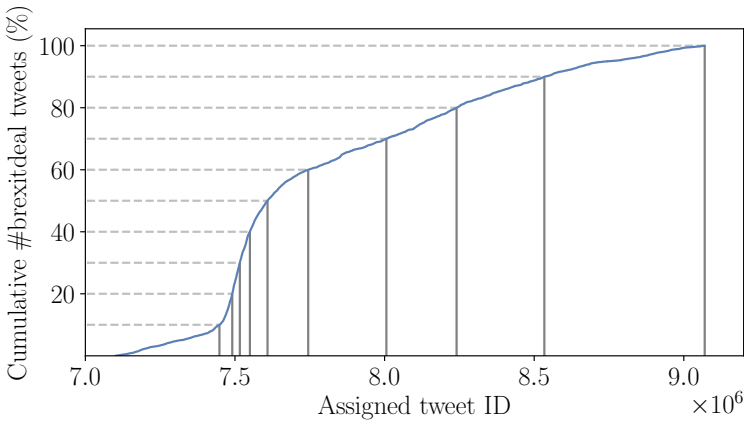
Fig. 7.4b. For this example, we choose $N = 10$. The dashed horizontal lines and solid vertical lines indicate the quantile values sampled (before normalisation). Finally the sampled IDs are normalised, differenced and inverted to obtain the activity shape shown in Fig. 7.4c.

Through the normalisation procedure and defining the final length of the representation, this methodology achieves our goal of facilitating fair comparison between different scales. The most important of these adjustments is the consideration of relative tweet rates rather than absolute counts since they are additionally unaffected by daily or weekly activity rhythms that can see notable changes in absolute counts under binning methods. One further advantage that this methodology presents is in the level of detail with which different periods of real time are viewed. Sampling from the CDF compresses periods of little activity and extracts more data points from periods of higher activity. This feature enables the representation to focus more on the main event period and is therefore less reliant on careful definition of the interval limits. This attribute is an important differentiating factor between our new approach and existing timeseries methods with N bins. Where the binning methods consider a series of time periods equally, our scale-independent representation automatically highlights event periods regardless of their timespan relative to the interval width.

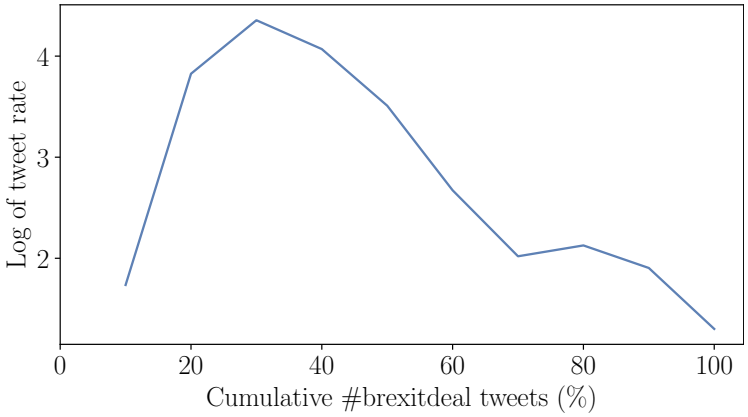
This method requires the choice of value for a single parameter, the number of quantiles. Our experiments tested $N = 25$, $N = 50$ or $N = 100$ quantiles (shown in Section 7.4) and found that $N = 50$ gave a good balance between smoothing the local noise and still retaining enough detail to characterise activity patterns. Provided that a sufficient minimum threshold is passed, we expect the number of quantiles to have a limited effect on the overall trends displayed. Intuitively, since we are counting the number of tweets required to see the next quantile of hashtag use, increasing the number of quantiles will reduce the component values in the final representation since we measure more intervals of fewer tweets, but the shape of the activity profile should remain similar.



(a) A portion of the hashtag's activity is considered, in this case around an increase in the relative usage rate.



(b) Quantiles in the CDF of dataset tweet IDs are calculated.



(c) The quantiles are used to measure the activity rate on the hashtag.

Figure 7.4: In order to resolve the differences in number of tweets between different timeseries segments we apply our CDF transformation. For reference, the daily binned timeseries is shown in Fig. 7.4a. In Fig. 7.4b we calculate the CDF and find the tweets observed between successive quantiles of interest. We use these tweet IDs to produce a vector of the desired length for each interval as shown in Fig. 7.4c. Here we set $N = 10$ for visual clarity, but use $N = 50$ for all subsequent analysis.

7.4 Results

In this section we demonstrate the utility of our scale-independent representation by looking at the intervals from a selection of 100 specimen hashtags. To ensure a

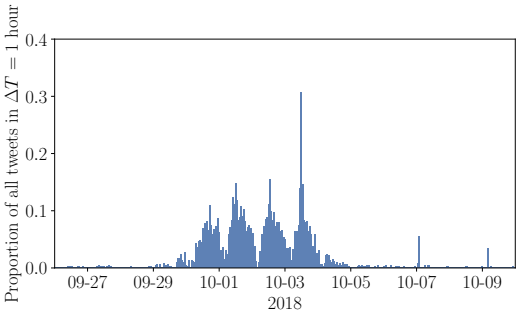
minimum level of "attention" in each interval and limit noise, we only consider those containing at least 100 tweets. This leaves 526 intervals for subsequent analysis, which can represent events or the lulls between them.

7.4.1 The importance of time resolution

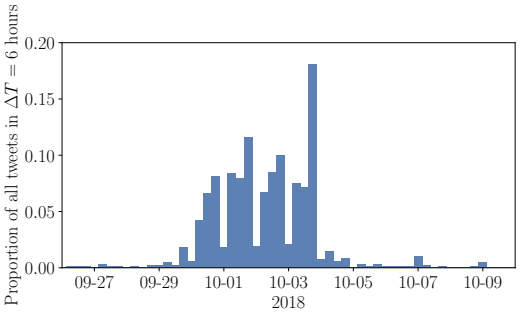
As a series of motivating examples, we first highlight how the appearance of the different classes proposed by Lehmann et al. are heavily influenced by the choice of bin size in the timeseries. In each of these examples, we use the methodology described by Lehmann et al. to identify the single largest peak usage of a hashtag. We then visualise it with hourly, six hourly and daily bins alongside the quantile-based hashtag usage rate profile representation proposed in this paper.

In Fig. 7.5 we explore *#cpc18*, a hashtag used at this time to show participation or interest in the 2018 Conservative party conference. The one-day width bins in Fig. 7.5c suggest that the period captured as an event by Lehmann et al.'s method is a single event and fits within their "before and during the peak" class. At the shorter timescales (Figs. 7.5a, 7.5b) it is revealed that the period actually covers a series of daily events peaking around the middle of the day as the events of the morning are discussed and anticipating the afternoon's news. The hashtag profile over this period reflects the daily repetition of smaller events and furthermore highlights the size difference between the activity on the peak and other days.

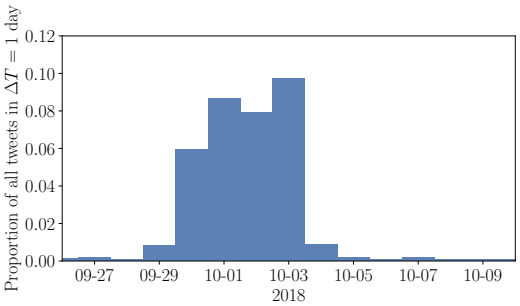
We next look at the peak-only class proposed by Lehmann et al. Our quick categorisation of this class only considers events for which the peak day accounts for 90% of usage activity for the hashtag. Fig. 7.6 shows one such hashtag *#esthermcvey*, in reference to the British MP Esther McVey who resigned as a cabinet minister on 15 November 2018 over opposition to the proposed Brexit deal. Social media attention to this event is short-lived, with most related tweets falling within a single day. With a smaller timescale it becomes apparent that this event is not characterised by a spontaneous end in the same way as its spontaneous start. The hourly resolution shows that interest tails off more gradually over a number of hours. The hashtag profile captures this behaviour, returning a right-tailed profile despite the majority



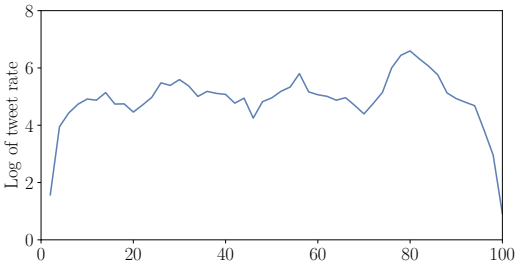
(a) Timeseries with one-hour bin width.



(b) Timeseries with six-hour bin width.



(c) Timeseries with one-day bin width.



(d) Scale-independent representation of the hashtag activity with $N = 50$.

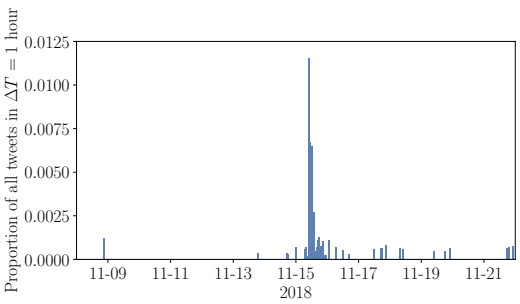
Figure 7.5: *#cpc18*, used to signify attendance or interest in the 2018 Conservative party conference. Choosing a bin width of one day presents this period as a single event, gradually building up to a peak. At shorter time resolutions, we see that activity rises and falls across each day of the conference. Fig. 7.5d shows the scale-independent representation of this period.

of the tweets occurring in a couple of hours.

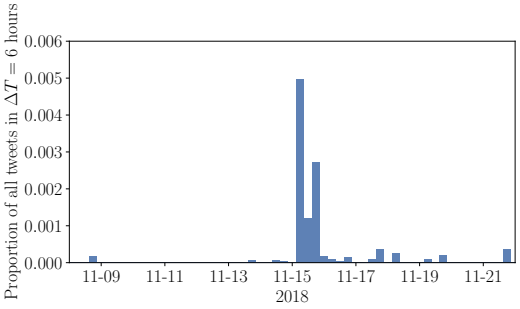
7.4.2 Characteristic shapes

We continue by examining the different profiles found under periods of increased hashtag usage rate to understand how the four classes of Lehmann et al. ("before and during the peak", "during and after the peak", "symmetrically around the peak" and "peak only" [111]) translate to our new methodology. We find four broad classes of profile, three of which match closely to three of the four classes proposed by Lehmann et al.

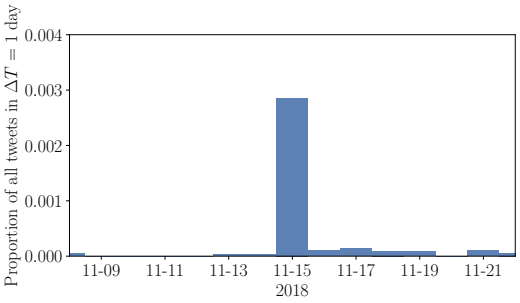
In Fig. 7.7a we present the first shape identified in periods of increased activity. This *right-tailed* profile is typical of unanticipated events such as announcements, interview comments or natural disasters. The hashtag usage rate (as defined in Section 7.3.3) increases abruptly and quickly peaks before gradually decreasing and returning to a background rate. Here we recover the "during and after the peak"



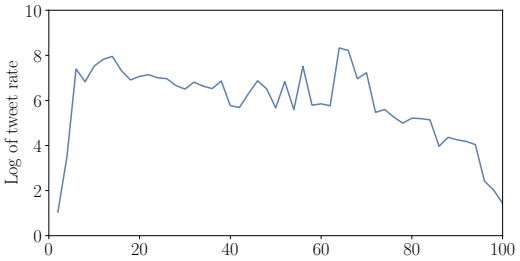
(a) Timeseries with one-hour bin width.



(b) Timeseries with six-hour bin width.



(c) Timeseries with one-day bin width.



(d) Scale-independent representation using the hashtag profile with $N = 50$.

Figure 7.6: *#esthermcvey*, referring to British MP Esther McVey and her resignation as a cabinet minister. Despite most of the use of this hashtag occurring on the peak day, there is evidence of a gradual decline on shorter timescales.

class from Lehmann et al.

In Fig. 7.7b we present an example of an *arch-shaped* profile characterised by a gradual increase of hashtag usage rate to the peak followed by an, often surprisingly symmetrical, decline to the background use rate. This profile shape corresponds to Lehmann et al.'s class "symmetrically around the peak" and constitutes many anticipated events which allow for both steady growth of attention, and sustained interest after the peak of activity, particularly scheduled events such as *confidencevote* and *peoplevotemarch*.

In Fig. 7.7c we provide an example of a *left-tailed* profile characterised by a gradual increase in hashtag usage rate to a peak, beyond which interest is not maintained, and usage rapidly returns to the background rate. This shape is the mirror complement of the right-tailed shape, common with events that have building anticipation beforehand but retain little relevance, or are quickly relabelled after the event. This shape most closely relates to the "before and during the peak" class observed by Lehmann et al.

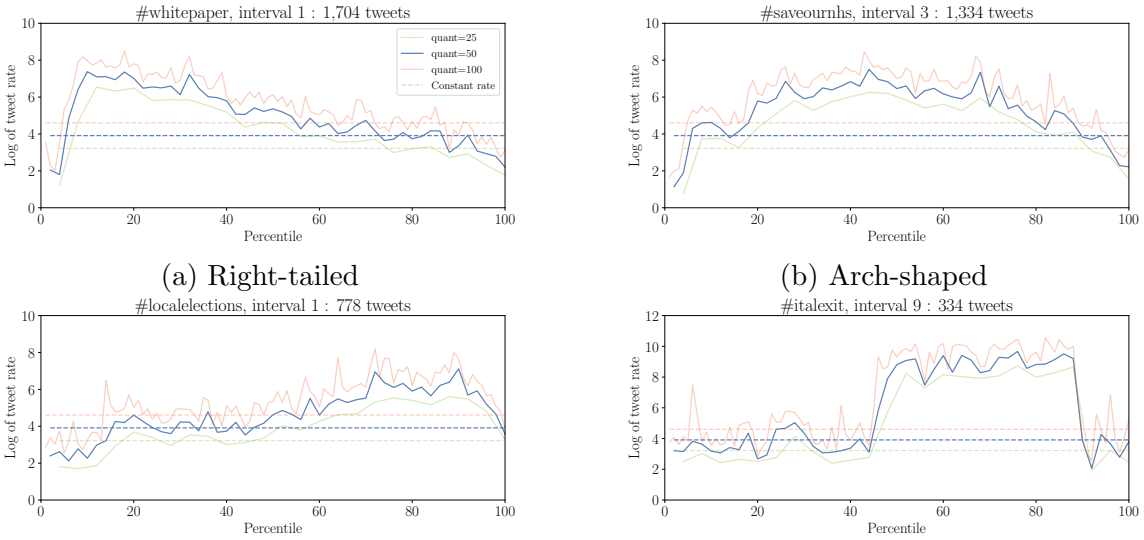
The final characteristic shape we observe is shown in Fig. 7.7d. Here we find

[...] Blow the whistle [...] #wikileaks #eagles #nfl #amazon #netflix #iTunes #Brexit #nyc #starbucks #uber #hiphop #music [...] lost even with DNC VoterFraud DOJ FBI Fake News Media Shadow Banning and Russia Dossier. Another Surprise waiting in 2018 ? [...]
#STOPCLANDESTINI #STOPINVASIONE #StopIslam #stopong #tolleranzazero #chiudiamoiporti #portichiusi #blocconavalesubito #iostoconsalvini #NessunoTocchiSalvini #italexit #frexit #grexit #nexit #brexit #stopEU #leaveEU #Stopsoros #SALVININONMOLLARE
The #Leave Campaign, aided by #Putin, used unfettered capitalism during the #BREXIT/#LEXIT Referendum @JohnMcDonnellMP [...]Why #Labour is @JeremyCorbyn helping @Conservatives to make Rupert Murdoch's day? #StopBrexit #WATON #ABTV #FBPE [...]
#Anonymous [...] #DAX #Money #Investing #Wealth #HNW #UHNW #Verm[...]gensverwaltung #roboadvisor #Brexit #NationalSiblingsDay [...]

Table 7.2: Selected examples of repeated tweet text during periods with an abrupt shift hashtag profile. For readability and anonymity, strings of unicode characters, URLS and private user mentions have been replaced with "[...]".

periods of greatly increased hashtag usage rates with sudden changes between a normal state and an event state. We call this an *abrupt shift* profile. This shape is distinctly different from the three other characteristic shapes shown in Fig. 7.7 by having no gradual change in rate on either side of the peak. We observed this behaviour in hashtags such as *amazon*, *italexit*, *putin* and *trump*. It is hard to intuit what collective interaction could lead to this type of hashtag usage rate profile. Manually studying the hashtags which create rate profiles of this type suggests commercial interests or spamming behaviour, heavily implying automated "bot" contributions to the dataset. Some examples of such tweets are shown in Table 7.2.

Lehmann et al. do not identify this as a distinct phenomenon, although in the case where the shift is of the order of one day in length, it would be miscategorised in their "peak only" class, that is an event with both an abrupt increase and decline around an event-driven peak. More generally, we find no clear indication of the "peak only" class being recovered under our representation. Certainly our dataset contains events in the above categories which fall within a 24-hour span, but our non-parametric approach has likely shown the underlying event behaviour to fall within the other three categories. This is not surprising since classification of events as "peak only" is highly dependent on the choice of bin width and we have seen that collective attention events lack a characteristic timescale, see e.g. Fig. 7.3.



(a) Right-tailed

(b) Arch-shaped

(c) Left-tailed

(d) Abrupt shift

Figure 7.7: The four characteristic shapes found around increased hashtag usage rate. Different coloured profiles indicate different numbers of quantiles in the hashtag profile (i.e. the length of the representation). The dashed line approximates the value of constant activity evenly distributed through the lifetime.

7.4.3 Discrete categories or a continuum?

The examples of the classes presented here are clear cut, but our results show that the hashtag usage rate profile produces a continuum of shapes rather than discrete classes as observed by Lehmann et al. Fig. 7.8 projects the vectors of hashtag profiles for different intervals into the two principal t-distributed stochastic neighbour embedding (t-SNE) components [180]. t-SNE is a conditional probability-based method for embedding higher dimensional points in lower dimensional space with additional optimisation for use on large datasets. The embedding aims to arrange points such that it maximises the likelihood of point x_i choosing other points as neighbours given a Student's t-distribution centred on x_i. There is a separation of left-tailed and right-tailed events, but rather than a distinct boundary, there is a continuous transition through arch-shaped profiles. Under this projection the abrupt shift class is not clearly distinguished from the three other classes.

In Fig. 7.9 we calculate the hashtag profiles to the four examples segments from Fig. 7.3. We can see that Figs. 7.9a, 7.9c, 7.9b and 7.9d all show the same right-tailed profile shape despite their differences in the binned timeseries. This example demonstrates the versatility of the hashtag profile to the choice of window captured.

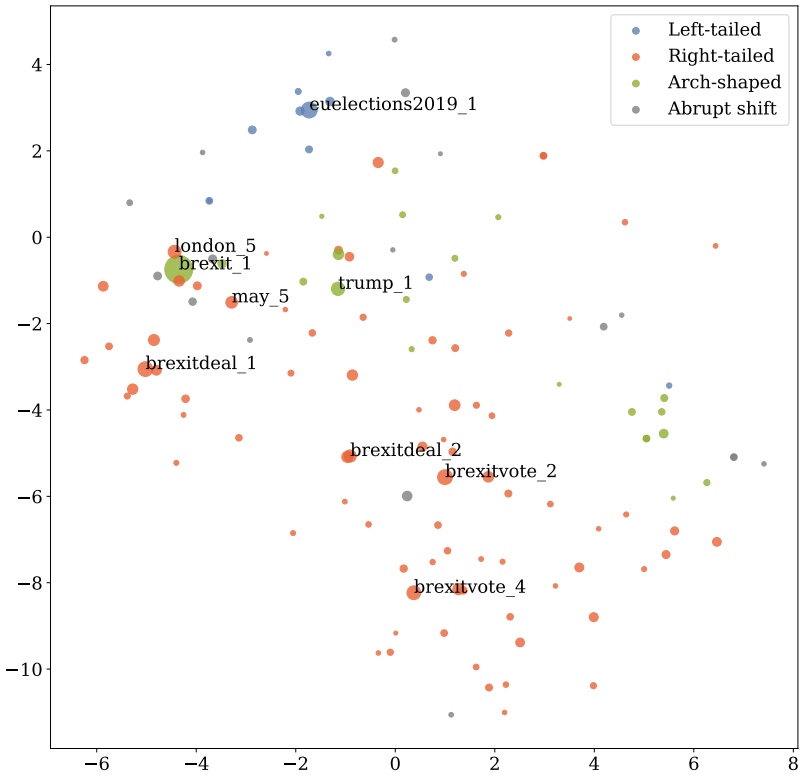
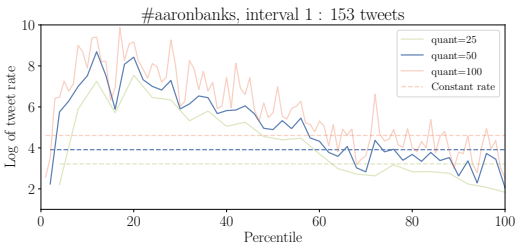


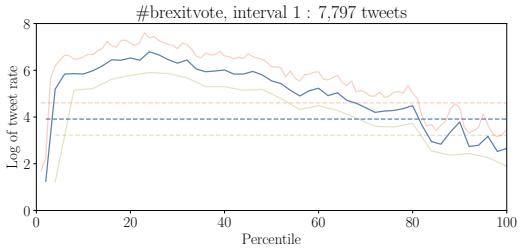
Figure 7.8: Projection of the scale-independent representations of hashtag intervals using the two dominant components under t-SNE analysis. Point colours denote the type classification by the authors and point size is proportional to the square root of total usage in the interval. Hashtag and interval number labels are provided for intervals with more than 10,000 tweets.

Each of the four examples here can be broadly characterised as right-tailed, but there are subtle differences between the shapes of the short-lived peaks and the gradual declines.

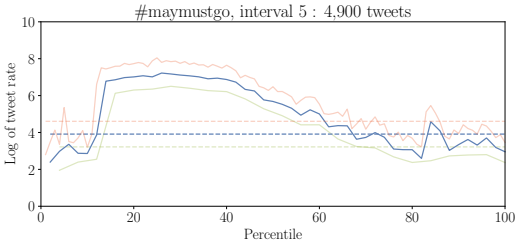
Of the periods identified in the hashtag usage, a total of 526 intervals of length at least 100 tweets were found. 78 were identified as right-tailed, 20 as arch-shaped, 10 as left-tailed, 16 as abrupt shifts, 32 as uniform (i.e. close to the level of constant use) and 370 did not clearly fall into one of the other categorisations. Identifying several intervals as uniform activity adds another useful attribute to the method. In these cases, the interval represents a period of normal activity between two events. With the ability to distinguish these periods, it is not necessary to identify event intervals during the splitting step, only the boundaries at which the activity profile changes. The fact that the majority of events do not clearly fall into one of the categories under visual inspection adds further support to the notion that these characteristic shapes are points on a continuum rather than discrete classes. Most of the intervals



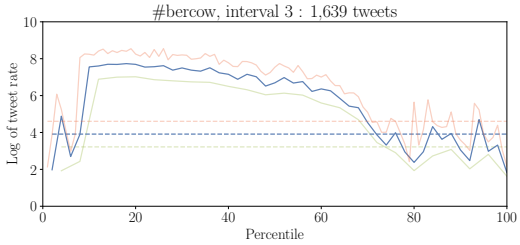
(a) Short-lived peak in days.



(b) Short-lived peak in hours.



(c) Gradual decline over days.



(d) Gradual decline over hours.

Figure 7.9: The hashtag usage rate profiles for the four examples from Fig. 7.3.

that were difficult to classify covered periods where the daily rate of hashtag usage oscillated rapidly, fell between two of the classes or contain multiple events that are clarified by the scale-independent representation. While this last situation arises from the limitations of the manual splitting process, it does demonstrate that our approach is still able to distinguish multiple events within a single interval.

7.5 Discussion

In this manuscript we presented a new method of studying the temporal trends of online activity, in particular how hashtag usage patterns change over periods of high interest. Our hashtag usage rate formulation accounts for two key limitations in large scale analysis of social media datasets by standardising both the total numbers of uses and the timeframe of interest, along with addressing the influence of exogenous drivers of activity such as time of day or longer-timescale trends. As a result, our method allows for a direct comparison of activity profiles across such scales that was difficult to manage before.

In Figs. 7.5 and 7.6 we show that the choice of timescales affects the apparent classes of different events and can obscure processes that occur over different lengths of time. Through these examples we show that the hashtag usage rate profiles represent the underlying dynamics of these hashtags and do not require any a priori

decision of which timescale to consider.

Our comparison with the four classes found by Lehmann et al. [111] shows that our method recovers the expected behaviour among common activity types. The "before and during the peak", "during and after the peak" and "symmetrically around the peak" classes from their work correspond to our left-tail, right-tail and arch-shaped profiles respectively, and demonstrate that they reflect an underlying process and are not an artefact of the choice of bin size when binning the timeseries. We do not find an analogue for the "peak only" class found by Lehmann et al. Intuitively, this is somewhat to be expected. Under Lehmann et al.'s approach certain short-lived events fall entirely within a single day and as such are reduced to a timeseries with a single non-zero data point. In this sense, the "peak only" category is operating as a filter category for event timescale. This is not an issue for our scale independent methodology which can be applied to events occurring over a single day, or many days, without any changes to the process. In addition, the time "dilation" effect granted by looking at tweet rate rather than tweet timestamps magnifies the single point relied upon by traditional methods and reveals that the underlying behaviour is often well-described by the other categories.

A special note should be made about the abrupt shift profile characterised by periods of time on hashtags such as *amazon*, *italexit*, *putin* and *trump*. This shape may look similar to Lehmann et al.'s "peak only" class but suggests more about the underlying system dynamics than the day aggregate counts. These events demonstrate none of the gradual build up or decline in attention levels suggestive of preferential attachment [17] that should be considered characteristic of true collective attention. They instead represent a change of mode between a baseline activity state and a more rapid event state. Such behaviour is unlikely to occur naturally in a self-organising system and suggests that certain exogenous perturbations are applied. After examining specimen tweets from these hashtags there is evidence that this shape is characteristic of astroturfing and other artificial attempts to increase visibility of certain content. In the case of hashtags like *amazon* and *italexit*, tweets during these periods of activity frequently include variations on a comment that is

trying to be publicised alongside a number of hashtags to promote visibility. This is a common tactic with malicious accounts, and will usually include several of the currently trending hashtags to promote visibility in as many searches as possible. The activation of automated accounts would also help to explain the suddenness of the transition between modes in the hashtag profile and the stark contrast from the more typical behaviour of more gradual transitions on one or both sides of the peak.

Following these observations, we should also highlight how the shapes observed in our scale-independent representation can be used to infer generating mechanisms more widely. Considering parts of the shape as either "steep" or "shallow" as determined by their gradient, we recall that steep changes are necessarily linked to periods where the activity rate of a given hashtag changes rapidly compared to the background rate. During the emergence of an event, such a steep period suggests some external stimulus is driving new attention, whereas a steep decline shows that the hashtag quickly loses relevance after its peak. Shallow increases, however, are more indicative of the preferential attachment-like behaviour required for our preferred definition of collective attention. Understanding the behaviour behind shallow decreases is more difficult, but such patterns likely arise from events which remain relevant beyond the initial peak through repeated updates or emergent conversations. These links with the underlying dynamics, and preferential attachment in particular, are an important advantage of our new method over existing timeseries techniques.

Despite our characterisation of four notable profiles, Fig. 7.8 shows that these shapes are not distinct categories but instead points on a continuum. In light of this we argue that the characteristic shapes we highlight above should be used to help infer the underlying behaviour of collective attention, rather than labelling a particular activity period. Tails in the hashtag profile suggest gradual changes in activity, whereas steep increases signify rapid changes. Understanding these features of the hashtag profiles may allow scholars to identify the mechanistic behaviour of the hashtag without reliance on categories which would be poorly fit to most of our dataset.

The main alternative approach for the analyses presented here would be using Lehmann et al.'s methodology with a variable bin size. Applying their method in this way has a number of disadvantages compared to our method, beyond the flaws in binning techniques mentioned already. Determining the appropriate bin width and number of bins to consider around the peak for each hashtag presents a similar manual challenge to dividing activity into intervals as required at present for our method. Lehmann et al.'s approach also specifically limits the density of events in time by requiring a minimum gap between peaks. While this requirement is not essential to the process, it will be difficult to distinguish aftershocks or secondary events from any natural rhythms in activity. Our consideration of rates accounts for this possibility and can still recognise multiple events in a single interval.

7.6 Future work

In light of the results in the preceding sections, we recommend that future analysis of behavioural trends in online social media should consider carefully whether binning at a certain timescale is the appropriate choice for the time period and media studied. Our hashtag profile methodology permits comparisons that are not possible under conventional approaches such as natural disasters in wildly different contexts.

Despite these successes, more work is required to make this method applicable on large datasets. Two algorithmic challenges remain unresolved: dividing timeseries into intervals and classifying intervals. As discussed in Section 7.3.2, experiments with peak detection and tipping point methods have not been able to identify appropriate interval boundaries such that the lifetime of a single event is captured within in a single interval. If such a technique can be found, many hashtag intervals can be efficiently constructed. Through this increased interval availability, it may be possible to identify themes which demonstrate the same activity profiles and are potentially subject to similar underlying processes driving the growth of attention.

Achieving this goal will need similar improvements to be made in interval activity classification methods. At present we consider the scale-independent representations of each interval as vectors and use vector techniques to investigate the relationships

between different intervals. As we saw in Fig. 7.8 these methods are broadly successful at differentiating the most visually distinct shapes but some cases prove more difficult. Imprecise trimming of the intervals also presents a challenge to vector based methods. Depending on the number of tweets before and after the event, it is possible for the activity profile of the interval to be shifted left or right through the components of the representation. This shift challenges vector-based methods and reduces the measured similarity when visual inspection would recognise the two shapes as the same. Having highlighted these necessary improvements, it should be noted that they are not entirely separate and solving one may alleviate the problems with the other. In this sense, the automated partitioning of the tweet timelines should be the preferred goal, enabling analysis at scale and likely improving the utility of vector-based similarity measures.

After an appropriate interval splitting method has been identified, the methods outlined in this manuscript can be applied to the dataset more widely, or other contexts. We believe the examples highlighted here are sufficient to demonstrate the merit of our methodology and the problems we address with it. The availability of a larger interval sample will facilitate an understanding of the relative frequency of the different activity profiles revealed by our method. It remains to be seen whether the comparatively high number of right-tailed intervals continues across the whole of the Brexit debate. As a frequent news topic during the study period, such a pattern would be plausible as breaking news captures widespread interest before decaying as the news cycle moves on. Comparisons with, and between, other topic areas such as sports, TV and movies may be similarly informative if attention to such topics can be shown to be driven by particular patterns of behaviour.

Chapter 8

Discussion

Throughout the previous four chapters we have seen how temporal trends in collective attention are reflected in online social networks. We defined the following set of research questions to motivate our investigations in this area.

- RQ1: How do structural patterns and polarisation in attention to information sources change under a disruptive event?
- RQ2: Which unipartite projection method allows community detection to best reflect ground truth in the bipartite network?
- RQ3: How do restrictions in offline spaces disrupt collective attention online?
- RQ4: Do repeated, similar events generate the same collective attention effects?
- RQ5: Can collective attention events be measured in a way that allows comparison across different scales?

In approaching this topic from a series of different methodological and experimental directions we have seen that extending the study of collective attention from static to dynamic perspectives facilitates a richer understanding of how group interactions and stimulating events influence the signals we see in social media data. We will now summarise the chapters in this thesis in the context of these research questions.

8.1 Is it correct to project and detect? How weighting unipartite projections influences community detection

Our exploration of collective attention first required the answer to a methodological question. In Chapter 4 we investigated a series of different weighting schemes for unipartite projection of bipartite networks, as previously published in [33, 34]. The seven candidate weighting schemes were drawn from those commonly used in various application domains, and include one designed to enable uniform contribution by each node in the network. A simple specimen network and a real-world dataset extracted from Twitter were then used to illustrate the wide variety of community structures detected by modularity maximisation when different projection weightings were applied to the same underlying network. Following this observation, we constructed a series of random bipartite networks with known community structure using several common degree distributions to test the accuracy of community detection under each of the seven candidate weighting schemes. We found that the performance of modularity-maximising community detection was better in the cases of degree distributions with longer tails. Across our selected evaluation metrics we found that the hyperbolic weighting scheme was preferable for capturing the nature of the bipartite community structure in the communities detected on the unipartite projection.

The experiments in Chapter 4 were designed to answer RQ2 and present a recommendation (and supporting evidence) for which approach should be preferred in future studies. Our choice of evaluation metrics extended the understanding of the interaction between the projection and community detection processes. We saw that four of the seven weighting schemes were susceptible to local noise creating large fluctuations in network characteristics in the low preference regime. Additional undesirable behaviour was similarly observed for several of the weighting schemes in terms of increased modularity when there was little underlying preference and either

too large or too small communities detected in the projection. As a result of these observations, our answer to RQ2 goes beyond advising which is most accurate to the ground truth by also highlighting cases where undesirable behaviour may occur. This additional insight will prove invaluable for future applications of this technique in experimental contexts where the ground truth is unknown.

8.2 Ideological biases in social sharing of online information about climate change

Having determined the preferred weighting scheme for unipartite projections of bipartite networks, Chapter 5 applied it to measure the impact of a disruptive news event on the online information ecosystem surrounding climate change. We saw how the announcement by then-US President Donald Trump of the withdrawal of the US from the Paris Agreement on climate change led to a massive increase in activity in the climate change discussion on Twitter. We used this extreme event to measure how the well-observed patterns of polarisation in climate change debates responded to such a large perturbation through the lens of URL sharing, as previously published in [35]. After collecting a dataset of tweets using the phrases *climate change* or *global warming* during the seven weeks centred on the Paris Agreement announcement, we constructed a series of bipartite networks for each week between Twitter users and the URLs they shared. These networks were subsequently projected onto the URL nodes where community detection found that the existing polarised trends in climate communications persisted, with URLs more likely to be shared alongside those with similar biases. More importantly, we found that the additional influx of users and sources during the peak of the event fit within the existing polarised structure without disrupting it over the course of the event. Considering the temporal aspect adds further strength to the findings of polarisation in online social networks by demonstrating that they are robust under disruptive events and are present outside of a committed core of individuals.

In reference to RQ1, Chapter 5 presents a potentially unexpected answer by in-

dicating that polarised structures are unaffected by a large disruptive event. The macro-level trends in the network structure are preserved and beyond this our focus on the giant component of each network means that any unobserved changes are necessarily separate from the main conversation. One likely explanation for these patterns lies in the self-selection of the users included in the study. Since we only consider those users invested enough in the debate to source and share information to their followers we should expect a level of investment in the climate change debate on their part. Alternatively, if there are true novices being drawn into the climate change debate, then the lack of disruption in the polarised structure suggests that media preferences may influence decisions on which content to share. These individuals may be new to the topic of climate change, but likely already have preferred news sources from which they gather their information. As reported by Schmidt et al. [164], such preferences generate polarised interaction networks. We should also consider what is required for a user to contribute to our projected URL networks. At a minimum, the user must share two different URLs in a given week. Therefore, the likelihood of contributing users being novices in the climate change debate may be small, even when the increased media attention draws new participants.

8.3 A tale of two lockdowns: Comparing mobility and attention responses in two periods of coronavirus restrictions in England and Wales

While the length of time passed during the Paris Agreement event was relatively short, other collective attention events can unfold over much longer timescales. Chapter 6 proceeded in this direction by considering attention to the COVID-19 pandemic and subsequent social distancing restrictions in England and Wales that evolved over the course of 2020. Using a dataset of geolocated tweets, we compared the two periods of most restrictive measures in the spring and autumn with equiva-

lent intervals from the previous year to better understand how offline restrictions are reflected in social media data. Attention to different locations was studied by dividing England and Wales into a grid of cells and comparing the relative frequency of tweets in each cell between 2020 and the same periods in 2019. Trajectories of Twitter users were calculated based on the distance and time between successive tweet locations, and again compared with the same periods in 2019. Analysis of spatial attention at the level of a month showed reductions in activity across much of the country and that locations away from populous urban centres were less frequently tweeted from under full lockdown conditions. During the early autumn, restrictions were more relaxed and this pattern is less apparent. Mobility of individuals also showed changes under lockdown conditions, as inter-tweet times and distances were biased towards shorter journeys compared with 2019. Once again, this difference was diminished during the autumn lockdown. Finally, we used a bipartite network representation of attention to topics on the days before and during each of the two full lockdowns. In the spring, topic structure was disrupted by breaking news stories around the pandemic and slowly returned to normal over the course of days and weeks. We observed in this way the greatest difference between responses to restrictions in the spring and autumn, as in the latter lockdown attention to the announcement was short-lived and less visible in the larger topic communities.

The three different perspectives of online social networks under pandemic conditions allow a rounded answer of RQ3 and how some manifestations of collective attention online are affected by offline restrictions. We saw that under the strictest levels of social distancing regulations, collective spatial attention was diminished in sparsely populated areas. In terms of structural changes at the topic level however, any disruption from the social distancing restrictions is short-lived before becoming incorporated into the existing patterns. In light of this, we have observed different responses in each of the two types of attention studied in Chapter 6. This is not completely surprising since the online conversations are not directly impacted by lockdowns, whereas the visitation of certain places (and their subsequent location tags) are necessarily curtailed. As a result of this observation, examination of re-

sponses to offline stimuli in online social networks should consider different contexts separately alongside their underlying relationship to the event.

By considering the two discrete periods of full lockdown restrictions in England and Wales during 2020 we have an excellent opportunity to address RQ4. We find additional support for the notion of "psychophysical numbing" identified by Dyer and Kolic [55] that expresses the concept of subsequent events needing to be more extreme to elicit the same emotional response. In comparison to the work by Dyer and Kolic, our work contrasts events with a much greater time gap which suggests that this numbing phenomenon persists in the longer term. We show that the impact on collective attention around the second lockdown is clearly diminished in terms of reach and permanence.

8.4 Shapes of collective attention online

The final contribution in this thesis presented a methodological improvement for studying temporal patterns in collective attention. Our new technique drew inspiration from the work of Lehmann et al. [111] who characterised collective attention events by the distribution of activity around their peak. We extended this concept to be scale-independent rather than rely on any predetermined temporal resolution. Chapter 7 details our new approach and demonstrates how the scale-independent nature of the new representation allows for direct comparison of events across vastly different time spans or total volume. We tested this new approach on a dataset of hashtags in the Brexit debate, each manually divided into a series of activity intervals. Using these example hashtags, we demonstrated that these representations visualise underlying characteristic shapes in the dynamics of collective attention to different topics that reflect three of the four classes Lehmann et al. identified. Additionally, we found that their "peak only" class was actually an artefact of the choice of temporal resolution, and since it depends entirely on sampling choices it should not be considered a true representation of attention dynamics. We found a fourth class under the scale-independent representation that suggests a new type of behaviour that was not detected in Lehmann et al.'s work. This behaviour sees

a rapid transition between two different attention states in the system, which thus far is suggestive of artificial attempts to manipulate attention to particular topics. Furthermore, our comparison of the scale-independent representation over multiple event periods implies that the characteristic shapes we identified represent a continuum of behaviours rather than discrete classes.

The examples shown in Chapter 7 demonstrate that the new method we propose enables the comparisons required across different time and volume scales to answer RQ5. This standardisation process is particularly important for identifying similarity in how attention to unrelated topics evolves during the course of an event. As an additional benefit, this technique can be applied to compare multiple events for the same topic to understand if attention dynamics are driven by repeated exposure to static processes or those that vary over time. As mentioned in Chapter 7, some more work is required before this technique can be applied for comparing a large number of events, but it is anticipated that such efforts will be well-rewarded for the opportunities they introduce for studying the dynamics of collective attention.

8.5 Critical reflection

The end result of this thesis has not advanced understanding of collective attention in precisely the way that was originally intended. Initially, the plan was to make use of the work of Borge-Holthoefer et al. [26] and use the structural changes in discussion networks observed around collective attention events to develop a predictive model for future events. Preliminary investigations into the feasibility of this approach raised a number of challenges. The first consideration is in the formulation of the predictive model. Borge-Holthoefer et al. required pruning of the network, which is suitable if the goal is to predict attention to a particular topic, but cannot be applied more generally. In such general situations, the networks become computationally difficult to handle, and the focus of attention becomes obscured among the many themes in the network. Other difficulties emerged in the choice of a representative null model and the suitability of such an approach for different event types, which led to the goals of this PhD thesis being adjusted towards better understanding the

structural and temporal patterns of collective attention.

After this change in direction, most of the work went as planned. One of the two difficulties of note arose when examining the bias of sources in Chapter 5. Obtaining objective ratings of biased opinions towards climate change as required for the characterisation of the URL communities in Chapter 5 proved impossible since, to our knowledge, such ratings are not available (unlike ratings on a political scale). As discussed previously, this necessitated the manual coding exercise.

The second challenge arose in defining the study area for Chapter 6. When designing the mobility experiments, it was decided that it was important to include a comparison with normal behaviour from previous years. Initially, it was intended to consider all of Europe, and ideally see changes tied to each countries relative pandemic conditions. Following adjustments to our Twitter collection processes late in 2019, a geographic collection was started to cover all of Europe. Unfortunately the collections running prior to the switch only covered England and Wales and necessitated this tighter geographic focus.

In some ways, the COVID-19 pandemic presented an opportunity for the work presented in this thesis. As it became clear that disruption from the pandemic was going to be a relevant event for months and across country borders, we were able to seize the opportunity to study collective attention over longer periods than available in any other recent event. Moreover, since it represents such a pervasive event in society there are likely to be very few aspects of social media conversations that have not been affected by the pandemic. These opportunities present a rare "silver lining" of the pandemic and suggest this period will be useful to scholars of collective attention for years to come.

With the benefit of hindsight, there are some aspects of this thesis that would have been approached differently. Revisiting the domain coding process from Chapter 5 would ideally correct the perceived discrepancies among some of the sites. Any means of doing this would require additional time commitments on the part of the graders, either through including more articles, more graders or larger text extracts. Altering the process for gathering the text extracts would likely be the best place

to start. Given the potential issues with extracts being taken out of context we discussed in Chapter 5, including more paragraphs taken consecutively would be the preferable next step while balancing the need for the sampled text to be anonymous with respect to its source.

The comparisons with years unaffected by the pandemic calculated in Chapter 6 are another area that could be revisited. Ideally, these comparisons would benefit from including more non-pandemic years as a benchmark. There is no reason to believe that 2019 was an atypical year for Twitter activity, and analysis at the level of one month still includes millions of tweets so there is no reason to doubt the validity of our results. That said, comparing with more non-pandemic years will help to understand any existing year-on-year changes in Twitter activity. The data required for such work were not available until early 2021, when Twitter released an academic API giving free access to historic data for research purposes. Unfortunately these data are not likely to be comparable with samples from the streaming API (e.g. by the exclusion of deleted tweets and accounts breaching Twitter's terms of service). Therefore to ensure a fair comparison, the existing dataset would also need to be replaced with one from the academic API.

8.6 Future work

After considering each of the contributions to the literature made in this thesis we have addressed each of the five research questions in detail. Through the course of these investigations however, a number of additional directions for future research have emerged.

There is extensive scope for extending the work studying the relationships between the choice of weighting scheme for unipartite projections of bipartite networks and subsequent network metrics. Allowing for different degree distributions in each of the bipartite modes is one such step. We were particularly interested in the case of social media networks where typically both modes have heavy-tailed zeta degree distributions, but other pairings of degree distributions exist in other contexts. Arthur [13] allows for such situations in his random network models, although only

a single projection weighting scheme is considered. One other direction in which to expand the work of Chapter 4 is in varying how the communities in each mode are related. This small change adds exponentially more cases to consider through the potential of different community sizes or numbers, necessarily breaking the mode symmetry in our existing random network model and requiring the left and right projections to be considered separately. Given this added complexity, addressing this gap is best done with a specific use case in mind.

Chapter 5 showed that the polarised structure of the climate change conversation was largely unaffected by the influence of a large collective attention event. An interesting avenue of future research would be to explore whether this trend exists in other polarised structures and other contexts. The causes and responses to the developing climate change crisis have become emotive subjects and as such very few neutral opinions are likely to exist. Coupled with this, expression of opinion on social media is likely to be self-selective of only the most invested individuals. Other political contexts may experience these same difficulties. Testing whether this polarised structure is universally resistant to disruption under collective attention events may require study of other contexts such as sporting events. Here we may be able to observe the polarised fans of two competing teams joined by neutral fans with no team affiliation. Understanding how these individuals fit within the existing team divide will complement the study of polarisation under the influence of collective attention events.

At the time of writing (spring 2021), the response to the COVID-19 pandemic in England and Wales has seen another extended period of lockdown across both countries. The comparison of the first two periods of full lockdown in Chapter 6 could be extended to include this additional period, plus any others that emerge over the course of the pandemic to give a greater exploration of the "psychophysical numbing" phenomenon noted by Dyer and Kolic [55] over a sustained period. Comparing these multiple lockdown periods raises an interesting question in terms of collective attention however. While the periods of lockdown may be distinct, the attention to the ongoing pandemic likely continues in the intervening periods of more relaxed

restrictions. Therefore more work may be required to understand whether these two lockdown events constitute discrete collective attention events, or are instead two moments in a single ongoing event. One final question that remains unanswered in terms of collective attention responses to the COVID-19 pandemic is how different countries, and their different regulations, respond to similar events in their own pandemic timelines. Such an approach may be particularly beneficial for identifying discrete events in the context of countries such as Australia and New Zealand who saw a return to a life largely without restrictions while the pandemic continued to spread in other areas. Given the dataset currently available, extending this work to other regions would require gathering new samples of historic tweets through the academic API as discussed previously.

The key question to emerge from the work in Chapter 7 is whether the characteristic shapes identified can be used to group topics based on similar underlying generating dynamics. We have already seen some evidence that this may be the case in the abrupt shift class that we define, which highlighted periods of spamming behaviour on the topics. To understand whether this applies more widely, and for the other topics, requires the development of suitable techniques for using this approach on large datasets without manual classification. One other avenue that may prove fruitful in advancing the understanding of attention dynamics is using this methodology in concert with agent-based models. Through simulation of information contagion in such settings it may be possible to link certain activity profiles with different meme characteristics. For example, it may be that seeding a meme with a high degree individual in the social network leads to a right-tailed shape. Understanding the relationship with other characteristics such as novelty, relevance and attractiveness will likely greatly improve the knowledge of how collective attention emerges.

8.7 Synthesis

Reviewing everything we've seen in this thesis, it is clear that the study of dynamic patterns of collective attention in online social networks provides important context

to how society responds to major events. Over the course of this manuscript we've shown that new insights can be drawn from considering existing social structures from a new perspective and pointed towards how these observations can be used to infer behaviour in the underlying system. Moreover we've shown that the study of collective attention is fertile ground for scholars who wish to investigate the interplay between real-world events and the evolution of social structures.

Social media data clearly has potential for studying modern society. This thesis has shown that the availability and scale of data make it a valuable tool for scholars, and can be used to understand communication patterns at the societal level. We've contributed to the knowledge of echo chambers in online communications and in particular how they respond to an influx of new participants, learning that while popularity may be fleeting in rapidly changing social media conversations, structural trends persist. More broadly, the links between offline events and online responses lend support to experimental studies and the merits of social media as a source of data and reinforces its relevance for future studies.

This thesis has also shown that collective attention is an integral part of how individuals communicate online. We've provided the foundations of a key link between the communication trends observed around different topics and their formation mechanisms. There is potential for this link to distinguish events comprised of groups independently paying attention to a topic to better understand situations arising from truly collective attention under our preferred definition. With further refinement, these advancements may form an effective means of identifying, and subsequently countering, malicious attempts to manipulate the attention of users of online social media.

Beyond the academic merits of its data, the role of social media in society is in a state of flux. Stories of attempts to influence political processes and stoke division in society are becoming more commonplace. Tied to this is the challenge of misinformation spreading through the ever-changing landscapes of social media, which has particularly hampered attempts to combat the COVID-19 pandemic. Regulatory measures have only recently begun to consider this aspect of social media

and it is unclear how these trends will ultimately affect current practices. What is clear however is that works like this thesis will help inform appropriate measures that limit the opportunities for malicious actors while still respecting the rights of individuals.

Social media and collective attention are still fields with much potential and many interactions with modern life. Many questions remain, and we hope that this work inspires and informs future scholars and their efforts to understand modern communication habits.

Appendix A

Supporting information to Ideological biases in social sharing of online information about climate change

Week	1	2	3	4	5	6	7
Users	13,017	16,553	22,496	54,347	19,033	17,574	14,252
URLs	6,117	6,560	7,871	20,880	8,529	7,652	7,272
Bipartite edges	17,990	22,565	31,831	80,009	25,165	23,701	19,702
Users in giant component	8,358	11,647	17,222	42,113	12,068	11,762	8,703
URLs in giant component	1,660	1,802	2,362	7,496	2,205	2,092	1,856
Bipartite edges in giant component	12,155	16,303	24,812	63,755	16,912	16,677	12,910
Unipartite edges	10,001	12,073	17,539	53,652	12,635	11,685	11,101
Modularity	0.606	0.564	0.549	0.578	0.613	0.566	0.595
Number of communities	44	37	57	117	44	45	44

Table A.1: Summary statistics for the networks across each of the seven weeks. In each case, the number of edges represents distinct edges and modularity is given to three decimal places.

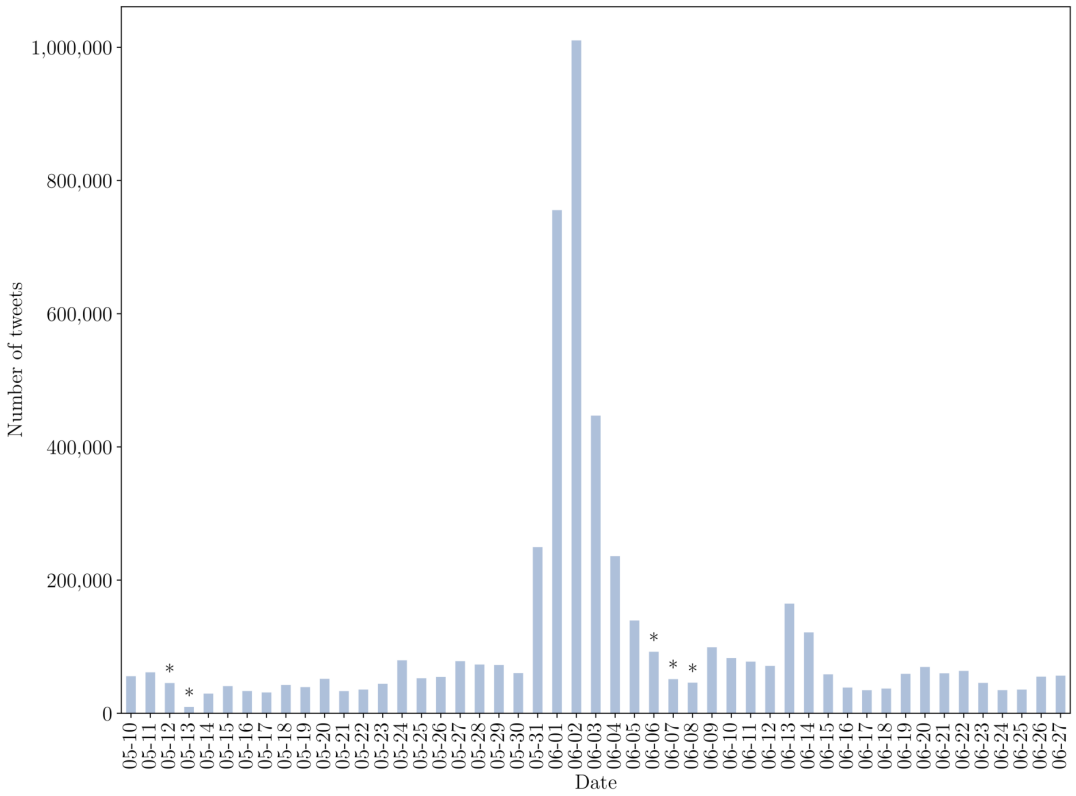


Figure A.1: Timeseries of the number of tweets per day across the seven week study period. There were a few short collection outages (11:00 2017-05-12 - 11:00 2017-05-13, 17:00 2017-06-06 - 09:00 2017-06-07 and 20:00 2017-06-07 - 10:00 2017-06-08) but since these outages represent a small proportion (around 5%) of the total collection period and mostly occurred at night, it is not expected that this has affected the validity of the experiments. Days with collection interruptions are marked with $*$. Note the increase in the number of tweets per day centered on 2017-06-02, and the evidence of weekly periodicity, particularly towards the end of June.

Week	Size of five largest communities				
1	423	264	148	123	93
2	326	271	254	212	127
3	528	354	303	209	198
4	1,732	1,263	954	762	629
5	524	391	237	201	140
6	448	350	347	139	119
7	347	264	257	200	160

Table A.2: Number of URL nodes in each of the five largest communities for each week in the study period.

Rating	-1	0	1	NaN
Political bias	L	N	R	U
Climate bias	E	N	S	U

Table A.3: Ideological coding scheme applied to each domain by the coders. This table uses the abbreviations L for Left, R for Right, E for Environmentalist, S for Sceptic, N for Neutral and U for Unclear.

theguardian.com 18,557/1,031	independent.co.uk 14,715/455	nytimes.com 12,603/535
washingtonpost.com 6,540/395	breitbart.com 4,172/205	twitter.com+ 3,949/1,364
youtube.com* 3,895/30	bbc(.co.uk or .com) 3,270/365	wordpress.com+ 2,544/1,719
dailycaller.com 2,535/228	insideclimatenews.org 2,345/157	wired.com 2,042/70
cbc.ca 1,831/232	foxnews.com 1,801/159	forbes.com 1,772/174
nationalgeographic.com 1,766/186	thehill.com 1,748/140	thegatewaypundit.com 1,699/44
bloomberg.com 1,633/193	thinkprogress.org 1,595/150	ecowatch.com 1,560/103
theconversation.com 1,526/115	politico.com 1,516/68	qz.com 1,373/87
theblaze.com 1,351/63	mashable.com 1,306/94	newslocker.com* 1,300/1,297
theatlantic.com 1,265/87	time.com 1,158/97	motherjones.com 1,135/110
ijr.com 1,132/28	scientificamerican.com 1,131/192	fastcompany.com 1,091/37
dailymail.co.uk 1,063/204	change.org 1,056/142	facebook.com+ 1,043/519
nasa.gov 1,032/99	zerohedge.com 1,030/52	futurism.com 1,026/83
naturalnews.com 984/67	infowars.com 916/46	telegraph.co.uk 913/150
cnn.com 913/197	newsweek.com 904/87	truthfeed.com 902/31
usatoday.com 892/96	foreignpolicy.com 844/36	iflscience.com 836/56
thetruthdivision.com 814/3	territoryairservices.com* 811/147	latimes.com 805/140
abc.net.au 791/145	rt.com 775/44	medium.com 767/223
buzzfeed.com 741/51	dailykos.com 710/95	prageru.com* 695/10
rightrelevance.com* 694/1	dailywire.com 661/50	reddit.com+ 657/393
reuters.com 642/208	fw.to* 600/162	trendolizer.com 598/595
newyorker.com 580/44	wattsupwiththat.com 568/138	weather.com 561/105
wsj.com 538/92	yournewswire.com 517/33	grist.org 514/111
vice.com 509/105	economist.com 508/34	investors.com 501/29
google.com* 494/56	paper.li* 489/144	weforum.org* 477/74

Table A.4: The 75 most common domains by share count across the seven weeks of the study period. These domains are ordered by total share count. Numbers indicate total number of shares/number of unique articles. * denotes the domains excluded from analysis due to incompatibility with the Diffbot API. + denotes the domains excluded as social media sites. Note that the number of unique URLs for youtube.com and google.com are artificially low due to the disambiguation step removing video identifiers.

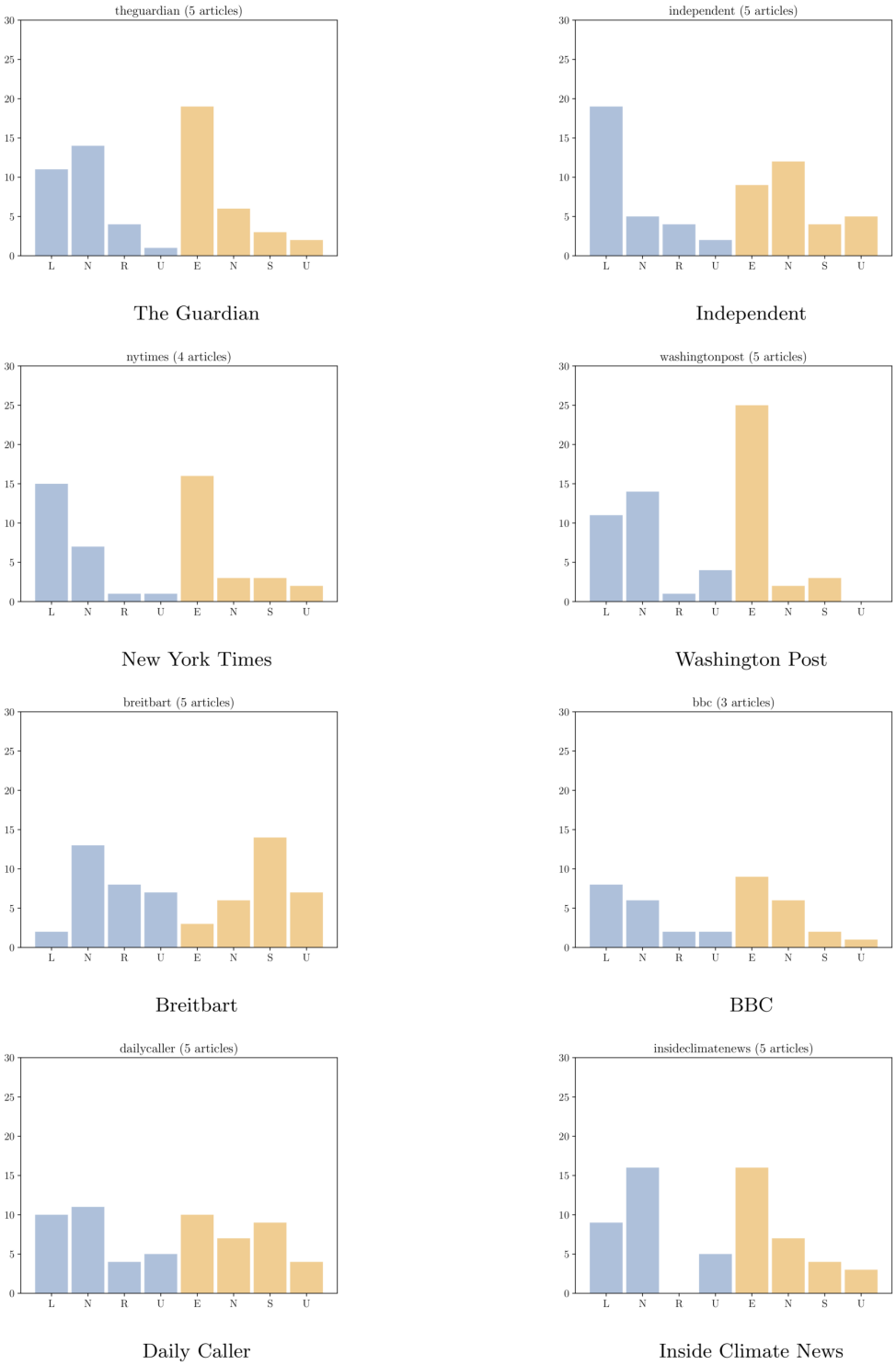
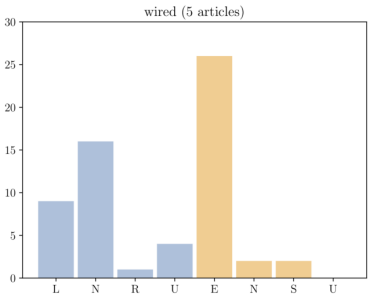
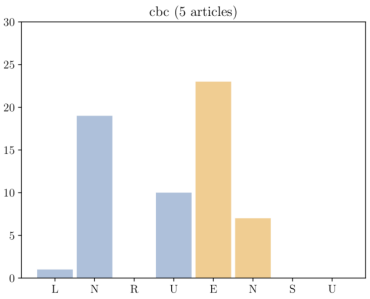


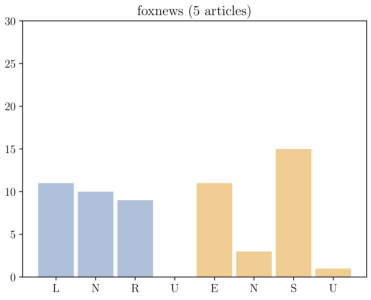
Figure A.2: Bias grades assigned by the coders to each domain.



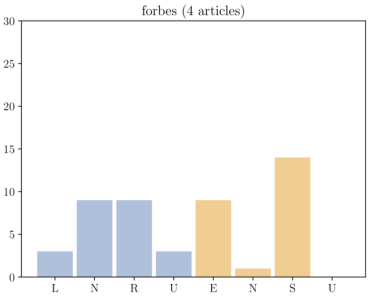
Wired



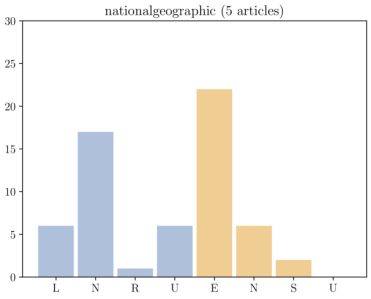
CBC



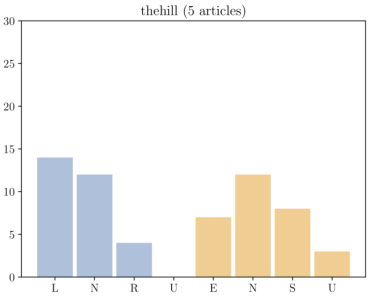
Fox News



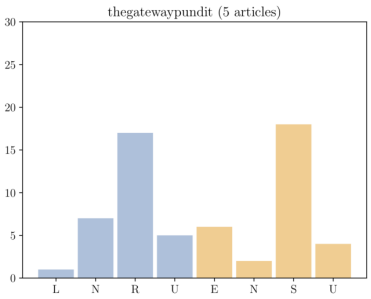
Forbes



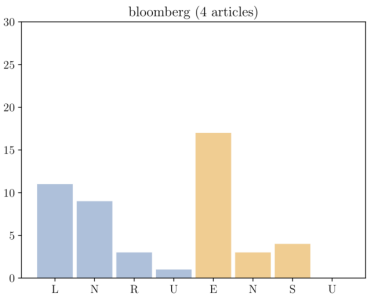
National Geographic



The Hill

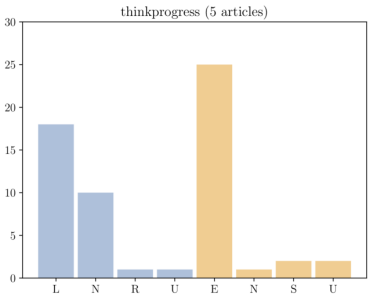


The Gateway Pundit

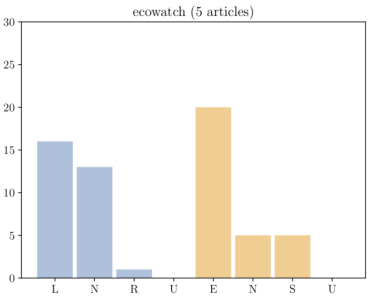


Bloomberg

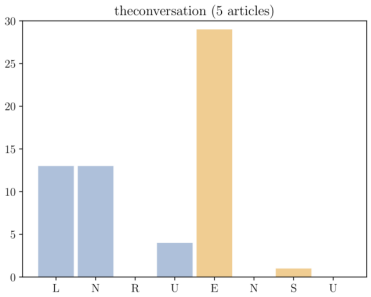
Figure A.2 (Cont.): Bias grades assigned by the coders to each domain.



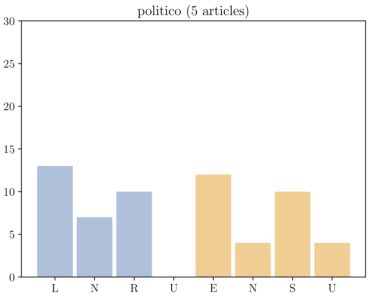
Think Progress



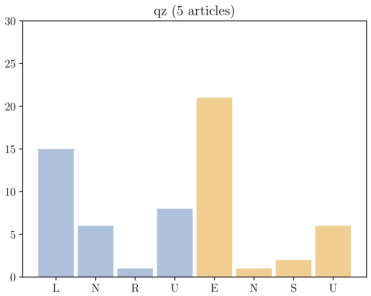
Ecowatch



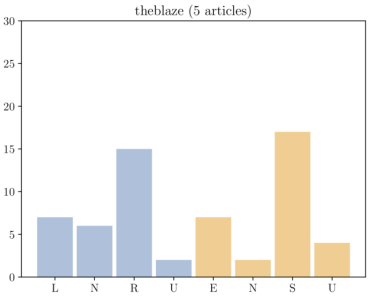
The Conversation



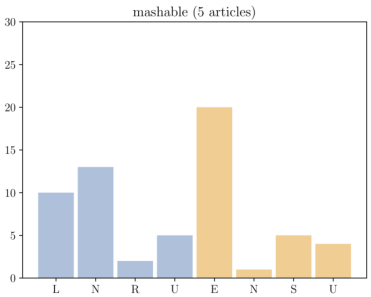
Politico



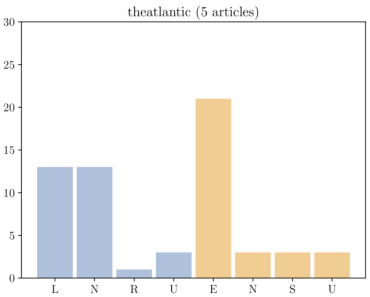
Qz



The Blaze

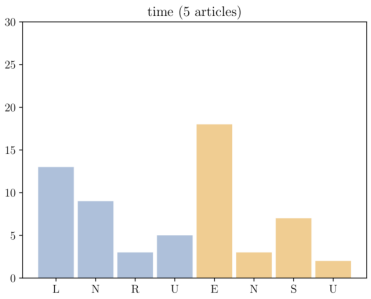


Mashable

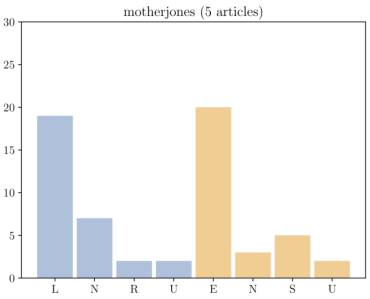


The Atlantic

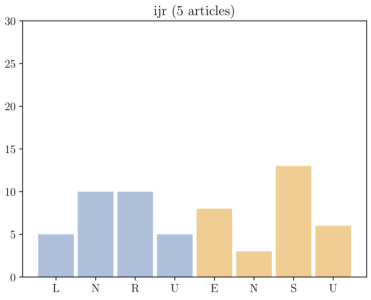
Figure A.2 (Cont.): Bias grades assigned by the coders to each domain.



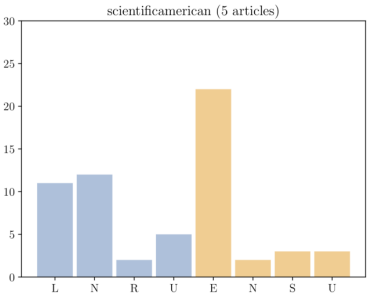
Time



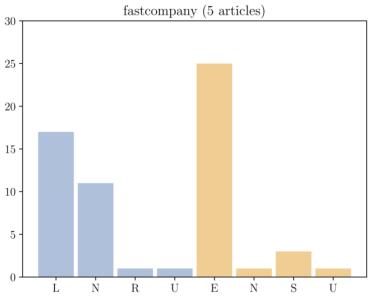
Mother Jones



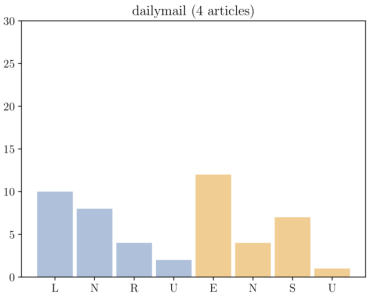
IJR



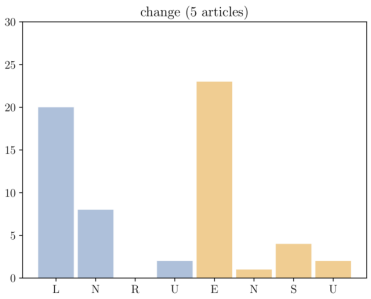
Scientific American



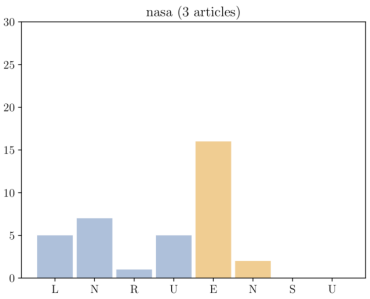
Fast Company



Daily Mail

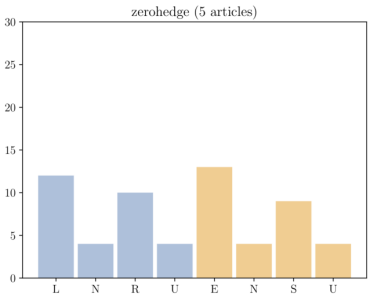


Change

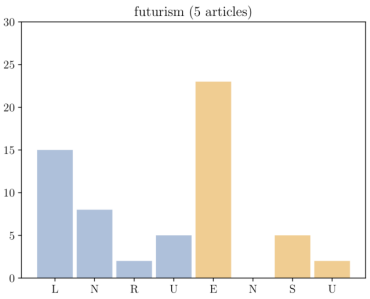


NASA

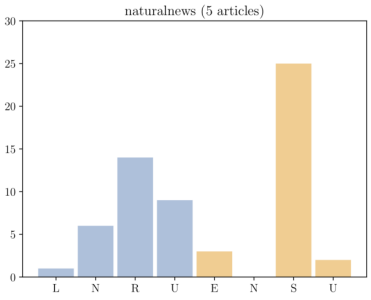
Figure A.2 (Cont.): Bias grades assigned by the coders to each domain.



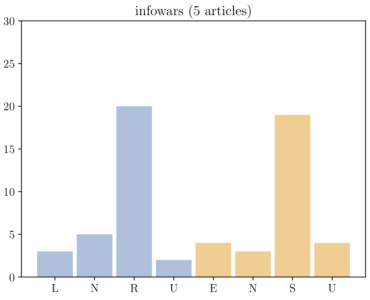
Zero Hedge



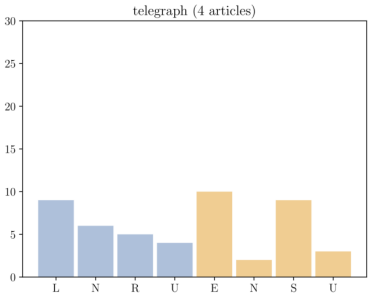
Futurism



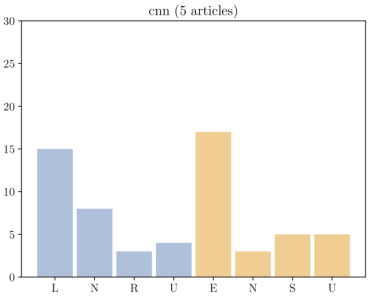
Natural News



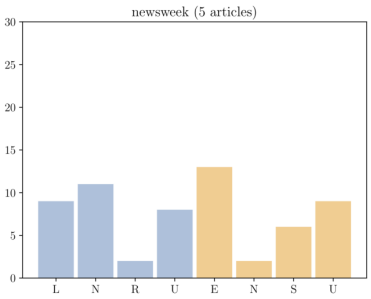
Info Wars



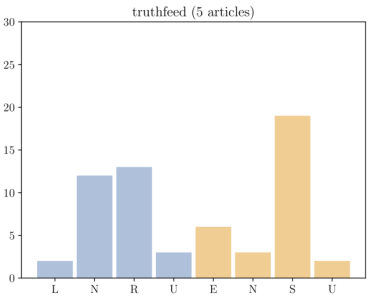
Telegraph



CNN

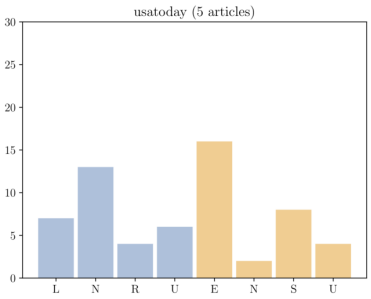


Newsweek

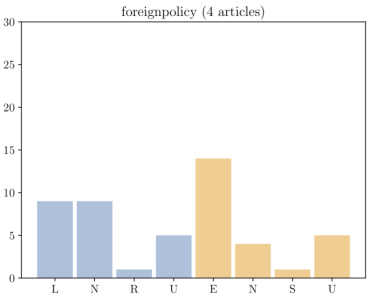


Truthfeed

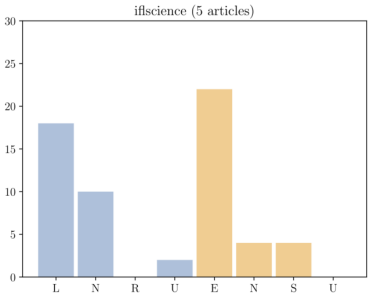
Figure A.2 (Cont.): Bias grades assigned by the coders to each domain.



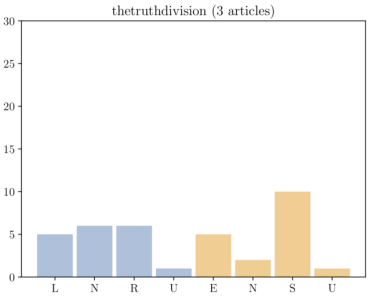
USA Today



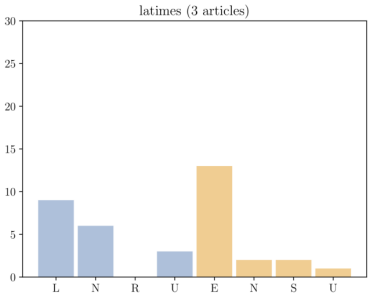
Foreign Policy



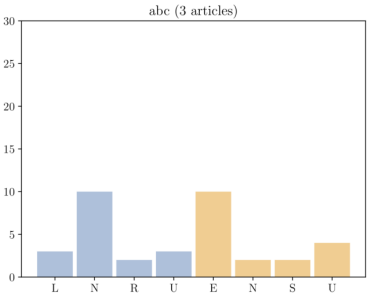
IFL Science



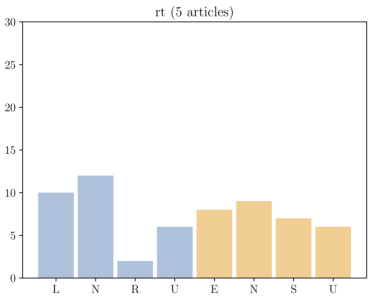
The Truth Division



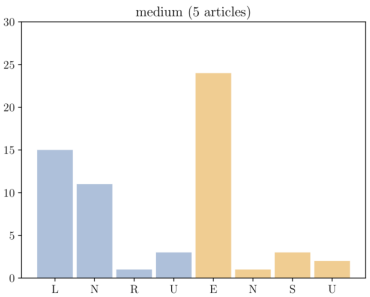
LA Times



ABC

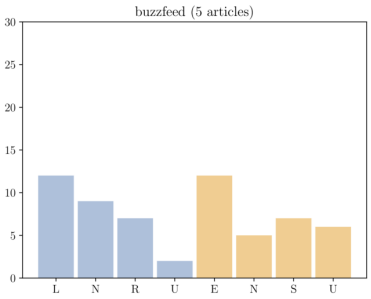


RT

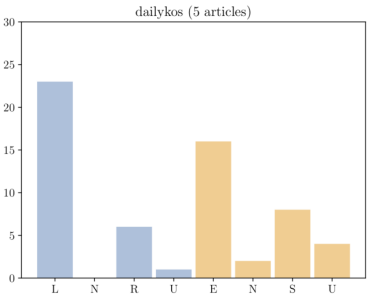


Medium

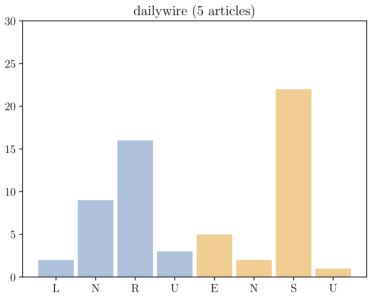
Figure A.2 (Cont.): Bias grades assigned by the coders to each domain.



Buzzfeed

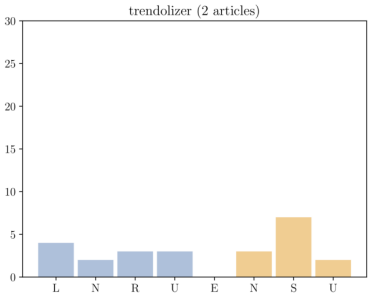


Daily Kos

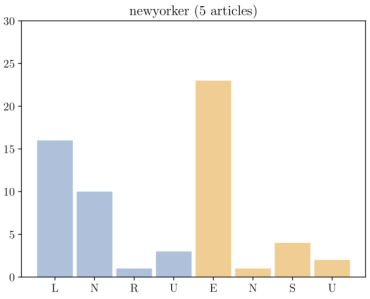


Daily Wire

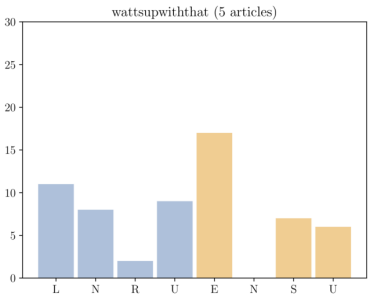
Reuters



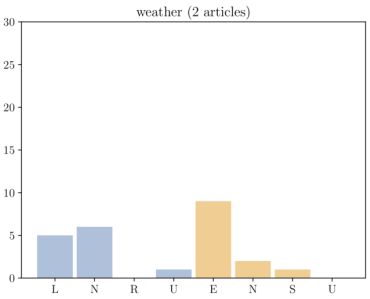
Trendolizer



New Yorker



Watts Up With That



Weather

Figure A.2 (Cont.): Bias grades assigned by the coders to each domain.

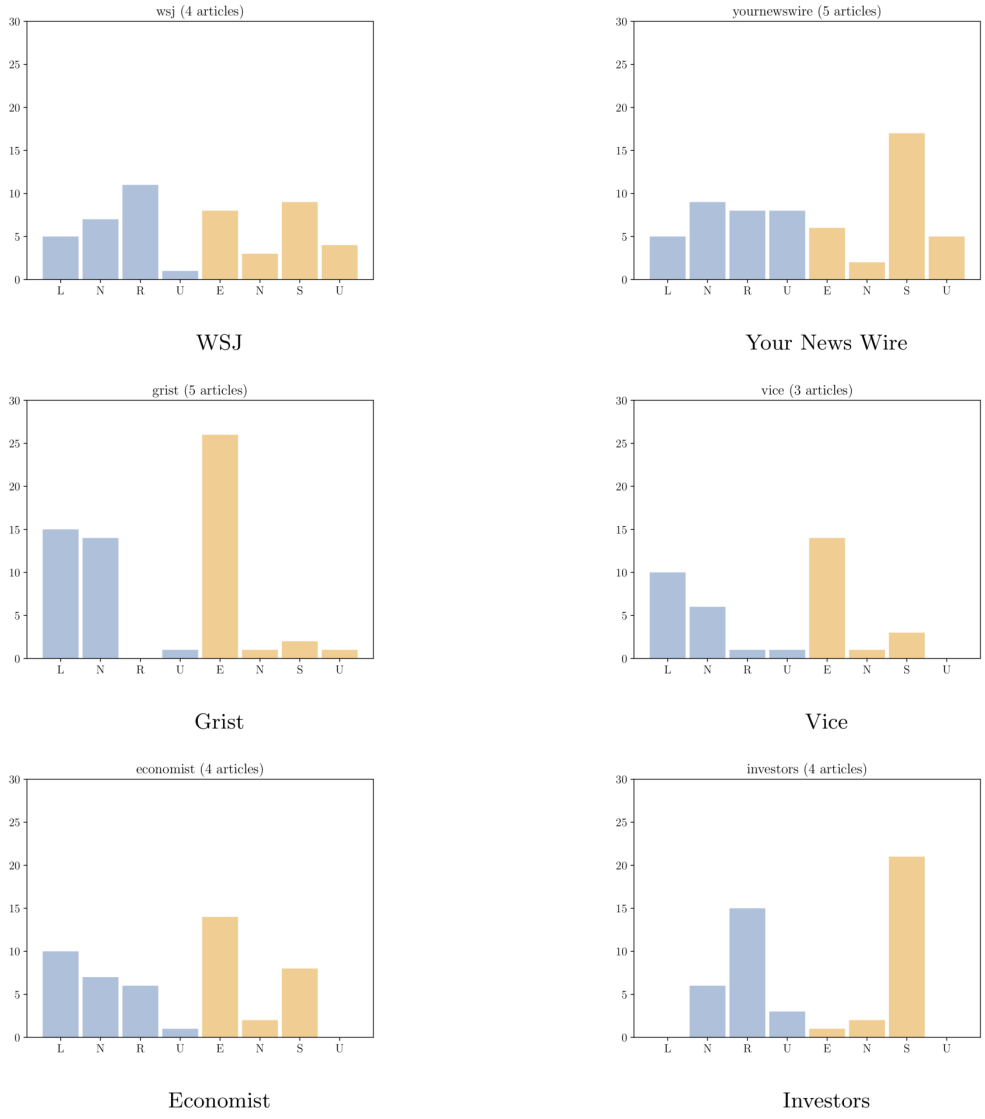
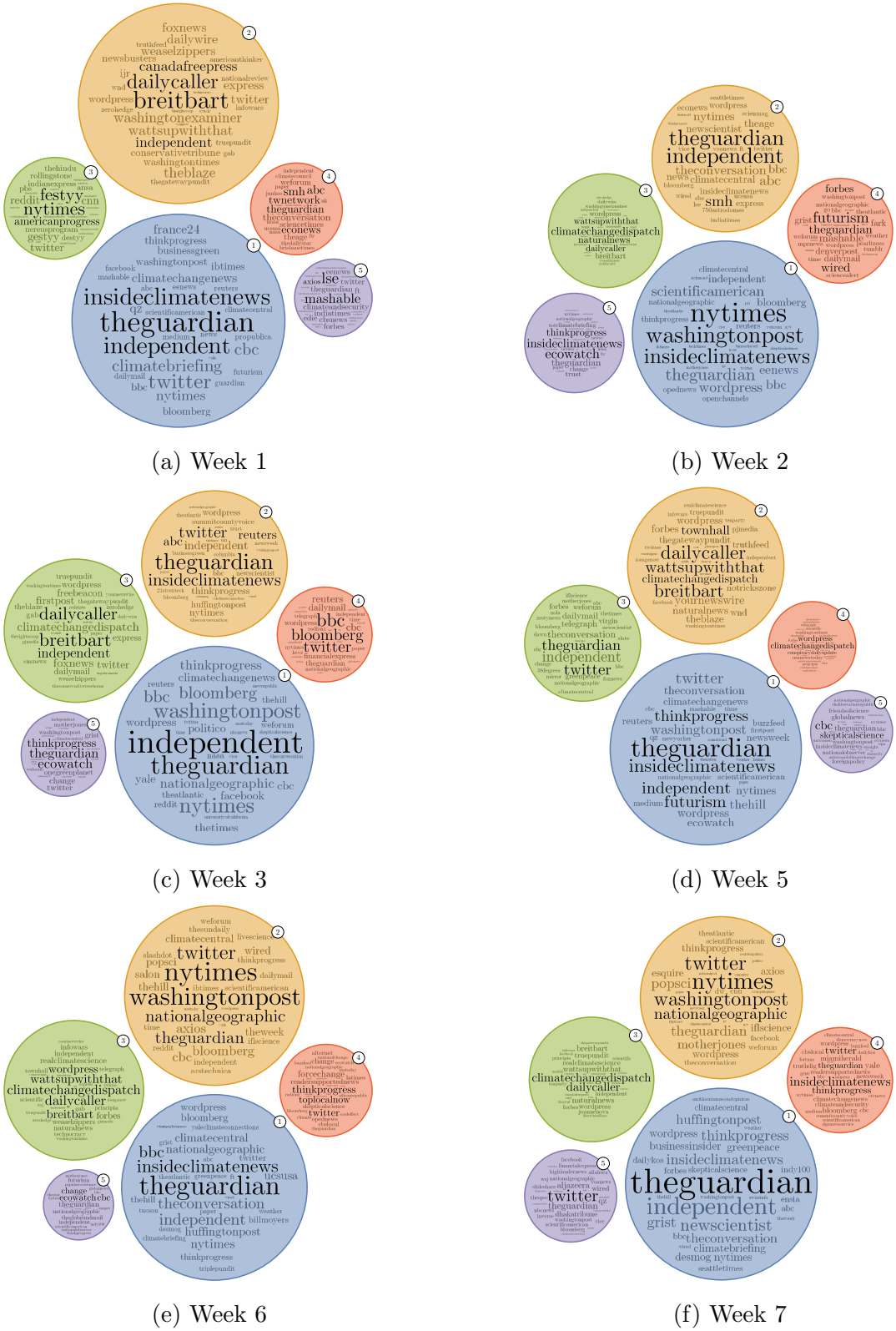


Figure A.2 (Cont.): Bias grades assigned by the coders to each domain.



(a) Week 1

(b) Week 2

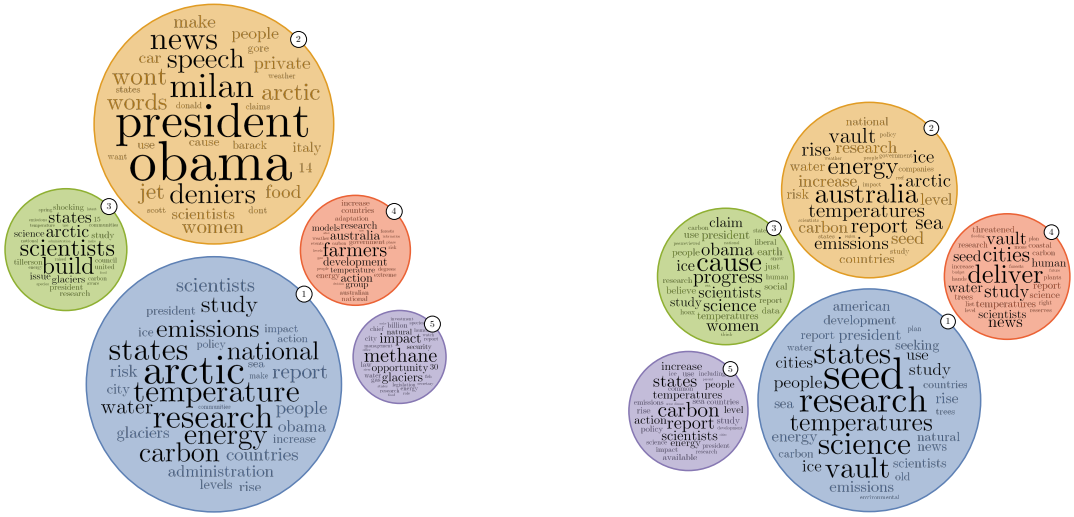
(c) Week 3

(d) Week 5

(e) Week 6

(f) Week 7

Figure A.3: TF-IDF weighted domain wordclouds for the five largest communities by share count over the remaining six weeks. Circle size is determined by the total number of shares for all URLs in the community. Terms coloured black are the highest weighted terms required to reach 15% of the total weight.

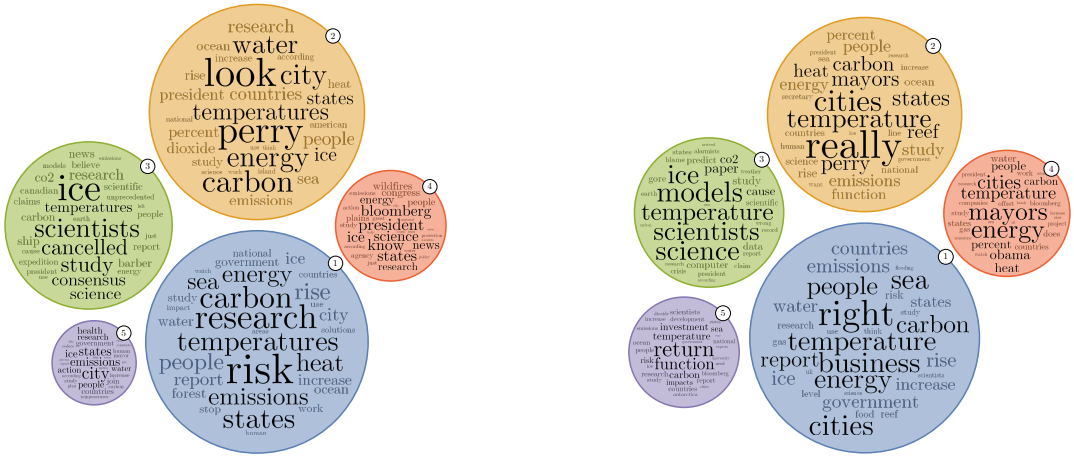


(a) Week 1

(b) Week 2

(c) Week 3

(d) Week 5



(e) Week 6

(f) Week 7

Figure A.4: TF-IDF weighted content wordclouds for the five largest communities by share count over the remaining six weeks. Circle size is determined by the total number of shares for all URLs in the community. Terms coloured black are the highest weighted terms required to reach 15% of the total weight.

(a) Week 1

(b) Week 2

(c) Week 3

(d) Week 5

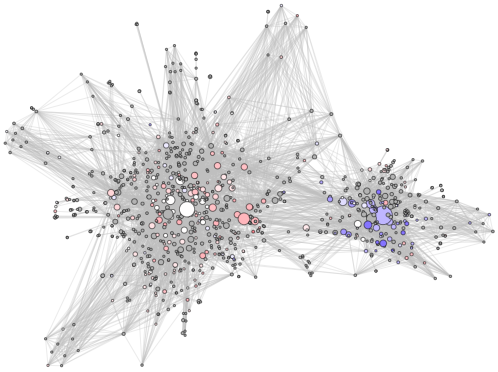
(e) Week 6

(f) Week 7

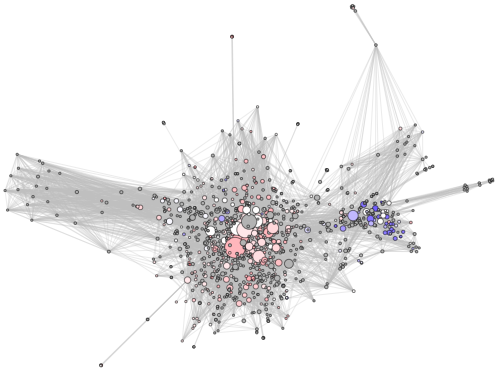
Figure A.5: TF-IDF weighted content bigram wordclouds for the five largest communities by share count over the remaining six weeks. Circle size is determined by the total number of shares for all URLs in the community. Terms coloured black are the highest weighted terms required to reach 15% of the total weight.

Persistence in the giant component

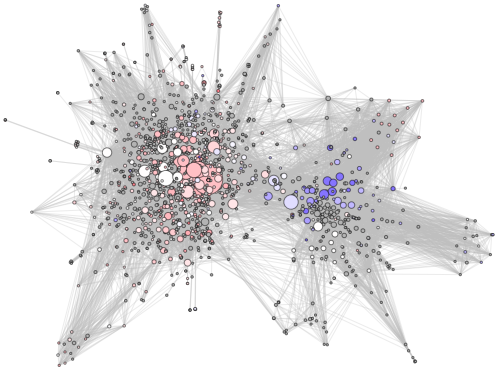
Restricting the analysis of persistence to the giant component in each week reduces the mean and variance but does not alter relative trends (user mean 0.231, user $\sigma =$ 0.084, URL mean 0.053, URL $\sigma = 0.049$, domain mean 0.486, domain $\sigma = 0.175$). Considering only the weeks surrounding Week 4 does not noticeably alter the mean, but does reduce the variance (user mean 0.236, user $\sigma = 0.039$, URL mean 0.076, URL $\sigma = 0.045$, domain mean 0.502, domain $\sigma = 0.074$.



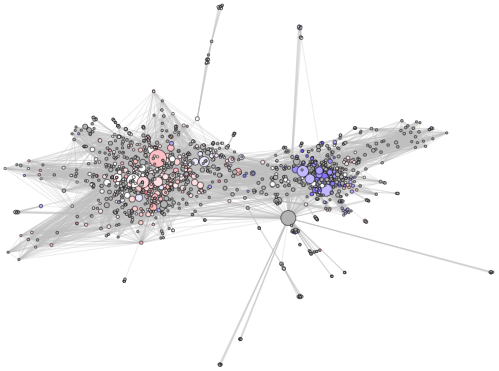
(a) Week 1: 60.24% of 1,660 nodes visible.



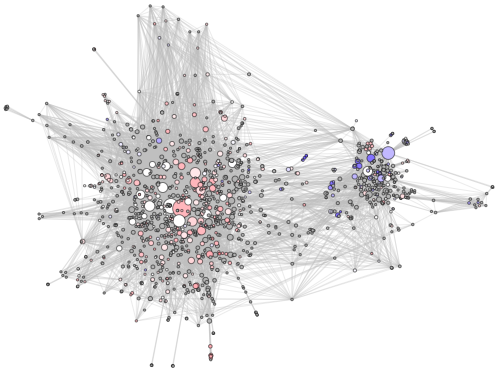
(b) Week 2: 66.04% of 1,802 nodes visible.



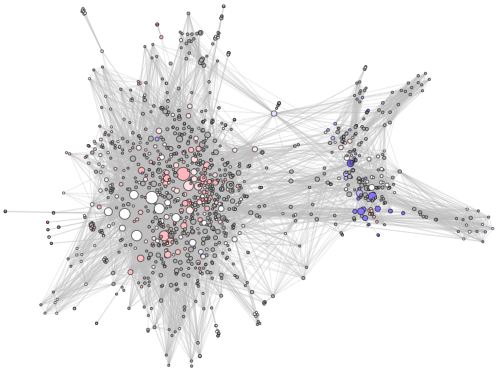
(c) Week 3: 67.40% of 2,362 nodes visible.



(d) Week 5: 67.71% of 2,205 nodes visible.

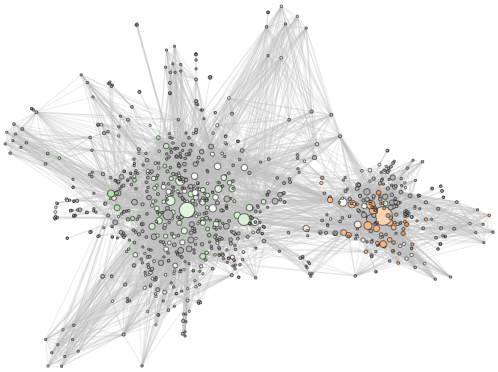


(e) Week 6: 67.07% of 2,092 nodes visible.

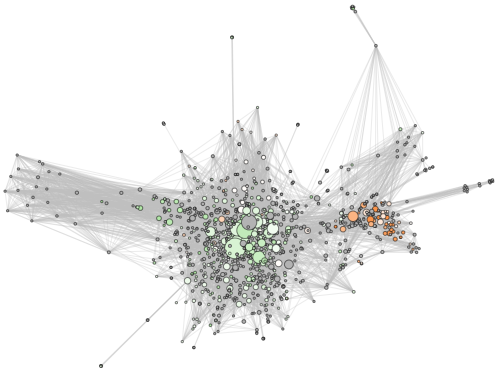


(f) Week 7: 66.16% of 1,856 nodes visible.

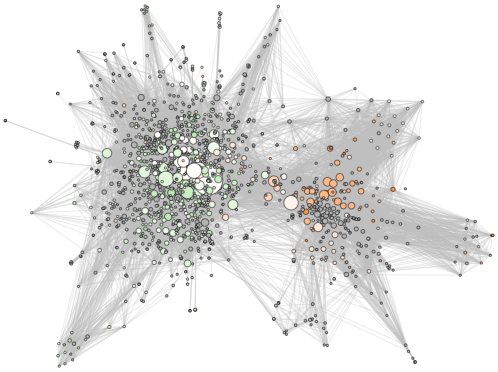
Figure A.6: Network diagrams of the top five communities across the six remaining weeks. Each figure is oriented such that the left-wing cluster is on the left and the right-wing cluster is on the right. In each case node colour signifies the political bias of domains as determined by the team of coders and size is proportional to the square root of total shares. Red nodes are from left-wing sources and blue nodes are from right-wing sources. Any node coded as neutral is white and grey indicates uncoded domains. Node placement is determined by the Python implementation of the ForceAtlas 2 algorithm [92]. The pattern of bias split between the clusters is consistent across the study period.



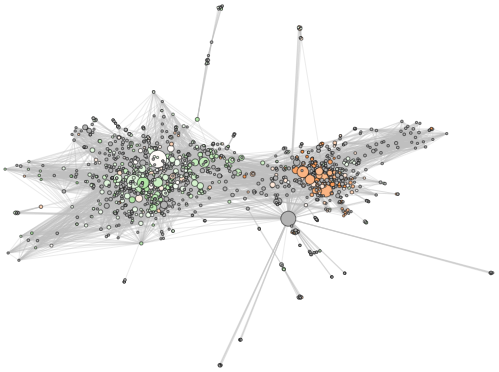
(a) Week 1: 60.24% of 1,660 nodes visible.



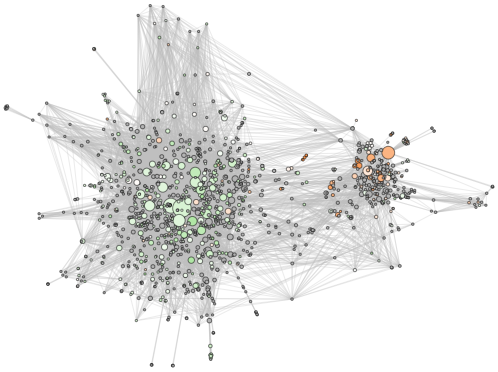
(b) Week 2: 66.04% of 1,802 nodes visible.



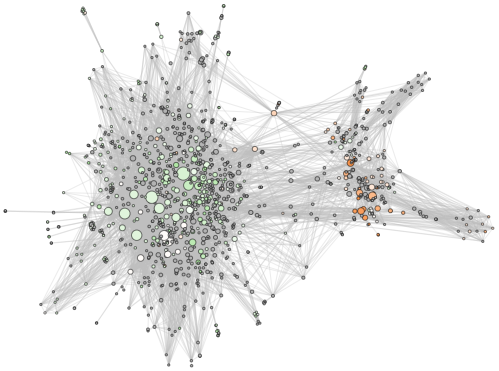
(c) Week 3: 67.40% of 2,362 nodes visible.



(d) Week 5: 67.71% of 2,205 nodes visible.



(e) Week 6: 67.07% of 2,092 nodes visible.



(f) Week 7: 66.16% of 1,856 nodes visible.

Figure A.7: Network diagrams of the top five communities across the six remaining weeks. Each figure is oriented such that the left-wing cluster is on the left and the right-wing cluster is on the right. In each case node colour signifies domain bias around climate change as determined by the team of coders and size is proportional to the square root of total shares. Green nodes are from environmentalist sources, orange nodes are from sceptic sources and any nodes from domains coded as neutral are white and grey indicates uncoded domains. Node placement is determined by the Python implementation of the ForceAtlas 2 algorithm [92]. Each week reveals the same pattern of polarisation in the network and demonstrates it is stable across the study period.

Bibliography

[1] Diffbot Analyze API developer documentation. https://www.diffbot.com/dev/docs/analyze/. Accessed: 2018-05-23.

[2] Media Bias/Fact Check. https://mediabiasfactcheck.com/. Accessed: 2018-07-06.

[3] Social Media Fact Sheet. https://www.pewresearch.org/internet/fact-sheet/social-media/. Accessed: 2020-11-30.

[4] Twitter Streaming API documentation. https://developer.twitter.com/en/docs/tweets/filter-realtime/api-reference/post-statuses-filter. Accessed: 2020-03-08.

[5] UNFCCC - What is the Paris Agreement? https://cop23.unfccc.int/process-and-meetings/the-paris-agreement/what-is-the-paris-agreement. Accessed: 2020-11-26.

[6] L. A. Adamic and N. Glance. The political blogosphere and the 2004 U.S. election: Divided they blog. In *Proceedings of the 3rd International Workshop on Link Discovery*, LinkKDD '05, pages 36–43. ACM, 2005.

[7] T. Agryzkov, P. Martí, L. Tortosa, and J. F. Vicent. Measuring urban activities using Foursquare data and network analysis: a case study of Murcia (Spain). *International Journal of Geographical Information Science*, 31(1):100–121, 2017.

[8] S. Ahajjam and H. Badir. Identification of influential spreaders in complex networks using HybridRank algorithm. *Scientific Reports*, 8(1):11932, 2018.

[9] M. Akbarpour and M. O. Jackson. Diffusion in networks and the virtue of burstiness. *Proceedings of the National Academy of Sciences*, 115(30):E6996–E7004, 2018.

[10] T. Alzahrani and K. J. Horadam. Analysis of two crime-related networks derived from bipartite social networks. In *Proceedings of the 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, ASONAM '14, pages 890–897, 2014.

[11] A. Arif, K. Shanahan, F.-J. Chou, Y. Dosouto, K. Starbird, and E. S. Spiro. How information snowballs: Exploring the role of exposure in online rumor propagation. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, pages 466–477, 2016.

[12] C. Armstrong, A. Poorthuis, M. Zook, D. Ruths, and T. Soehl. Challenges when identifying migration from geo-located Twitter data. *EPJ Data Science*, 10(1):1, 2021.

[13] R. Arthur. Modularity and projection of bipartite networks. *Physica A: Statistical Mechanics and its Applications*, 549:124341, 2020.

[14] D. Badre. How we can deal with "pandemic fatigue". `https://www.scientificamerican.com/article/how-we-can-deal-with-pandemic-fatigue/`. Accessed: 2021-03-17.

[15] E. Bakshy, S. Messing, and L. A. Adamic. Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6239):1130–1132, 2015.

[16] D. Balcan, V. Colizza, B. Gonçalves, H. Hu, J. J. Ramasco, and A. Vespignani. Multiscale mobility networks and the spatial spreading of infectious diseases. *Proceedings of the National Academy of Sciences*, 106(51):21484–21489, 2009.

[17] A.-L. Barabási. The origin of bursts and heavy tails in human dynamics. *Nature*, 435(7039):207–211, 2005.

[18] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

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

[20] A.-L. Barabási, R. Albert, and H. Jeong. Mean-field theory for scale-free random networks. *Physica A: Statistical Mechanics and its Applications*, 272(1):173–187, 1999.

[21] M. J. Barber. Modularity and community detection in bipartite networks. *Phys. Rev. E*, 76:066102, 2007.

[22] K. C. Bathina, A. Jammalamadaka, J. Xu, and T.-C. Lu. An agent-based model of posting behavior during times of societal unrest. In *Social, Cultural, and Behavioral Modeling*, pages 53–59, 2017.

[23] S. J. Beckett. Improved community detection in weighted bipartite networks. *Royal Society Open Science*, 3(1), 2016.

[24] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008, 2008.

[25] C. Bongiorno, A. London, S. Miccichè, and R. N. Mantegna. Core of communities in bipartite networks. *Phys. Rev. E*, 96:022321, 2017.

[26] J. Borge-Holthoefer, R. A. Baños, C. Gracia-Lázaro, and Y. Moreno. Emergence of consensus as a modular-to-nested transition in communication dynamics. *Scientific Reports*, 7(1):41673, 2017.

[27] J. Borge-Holthoefer, W. Magdy, K. Darwish, and I. Weber. Content and network dynamics behind Egyptian political polarization on Twitter. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, CSCW '15, page 700–711, 2015.

[28] C. Boulton and T. Lenton. A new method for detecting abrupt shifts in time series. *F1000Research*, 8(746), 2019.

[29] J. Bright. Explaining the Emergence of Political Fragmentation on Social Media: The Role of Ideology and Extremism. *Journal of Computer-Mediated Communication*, 23(1):17–33, 2018.

[30] J. J. Brown and P. H. Reingen. Social ties and word-of-mouth referral behavior. *Journal of Consumer Research*, 14(3):350–362, 1987.

[31] P. Burnap, M. L. Williams, L. Sloan, O. Rana, W. Housley, A. Edwards, V. Knight, R. Procter, and A. Voss. Tweeting the terror: modelling the social media reaction to the Woolwich terrorist attack. *Social Network Analysis and Mining*, 4(1):206, 2014.

[32] D. S. Callaway, J. E. Hopcroft, J. M. Kleinberg, M. E. J. Newman, and S. H. Strogatz. Are randomly grown graphs really random? *Phys. Rev. E*, 64:041902, 2001.

[33] T. J. B. Cann, I. S. Weaver, and H. T. P. Williams. Is it correct to project and detect? Assessing performance of community detection on unipartite projections of bipartite networks. In *Complex Networks and Their Applications VII*, pages 267–279, 2019.

[34] T. J. B. Cann, I. S. Weaver, and H. T. P. Williams. Is it correct to project and detect? How weighting unipartite projections influences community detection. *Network Science*, page 1–19, 2020.

[35] T. J. B. Cann, I. S. Weaver, and H. T. P. Williams. Ideological biases in social sharing of online information about climate change. *PLOS ONE*, 16(4):1–25, 2021.

[36] M. Castells. Communication, power and counter-power in the network society. *International Journal of Communication*, 1(1), 2007.

[37] C. Castillo, M. El-Haddad, J. Pfeffer, and M. Stempeck. Characterizing the life cycle of online news stories using social media reactions. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, CSCW '14, page 211–223, 2014.

[38] D. Centola. The spread of behavior in an online social network experiment. *Science*, 329(5996):1194–1197, 2010.

[39] S. Chang, E. Pierson, P. W. Koh, J. Gerardin, B. Redbird, D. Grusky, and J. Leskovec. Mobility network models of COVID-19 explain inequities and inform reopening. *Nature*, 589(7840):82–87, 2021.

[40] Y.-Z. Chen, N. Li, and D.-R. He. A study on some urban bus transport networks. *Physica A: Statistical Mechanics and its Applications*, 376:747 – 754, 2007.

[41] J. Choi and J. K. Lee. Investigating the effects of news sharing and political interest on social media network heterogeneity. *Computers in Human Behavior*, 44:258 – 266, 2015.

[42] A. Clauset, M. E. J. Newman, and C. Moore. Finding community structure in very large networks. *Phys. Rev. E*, 70:066111, 2004.

[43] E. Colleoni, A. Rozza, and A. Arvidsson. Echo chamber or public sphere? Predicting political orientation and measuring political homophily in Twitter using big data. *Journal of Communication*, 64(2):317–332, 2014.

[44] P. Connor, E. Harris, S. Guy, J. Fernando, D. B. Shank, T. Kurz, P. G. Bain, and Y. Kashima. Interpersonal communication about climate change: how messages change when communicated through simulated online social networks. *Climatic Change*, 136(3):463–476, 2016.

[45] M. Conover, J. Ratkiewicz, M. Francisco, B. Gonçalves, F. Menczer, and A. Flammini. Political polarization on Twitter. In *International AAAI Conference on Web and Social Media*, pages 89–96, 2011.

[46] M. D. Conover, B. Gonçalves, A. Flammini, and F. Menczer. Partisan asymmetries in online political activity. *EPJ Data Science*, 1(1):6, 2012.

[47] W. Cota, S. C. Ferreira, R. Pastor-Satorras, and M. Starnini. Quantifying echo chamber effects in information spreading over political communication networks. *EPJ Data Science*, 8(1):35, 2019.

[48] S. Cowie, R. Arthur, and H. Williams. @choo: Tracking pollen and hayfever in the UK using social media. *Sensors*, 18(12):4434, 2018.

[49] M. De Domenico and E. G. Altmann. Unraveling the origin of social bursts in collective attention. *Scientific Reports*, 10(1):4629, 2020.

[50] M. Del Vicario, A. Bessi, F. Zollo, F. Petroni, A. Scala, G. Caldarelli, H. E. Stanley, and W. Quattrociocchi. The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, 113(3):554–559, 2016.

[51] M. Del Vicario, A. Scala, G. Caldarelli, H. E. Stanley, and W. Quattrociocchi. Modeling confirmation bias and polarization. *Scientific Reports*, 7(1):40391, 2017.

[52] M. Del Vicario, F. Zollo, G. Caldarelli, A. Scala, and W. Quattrociocchi. Mapping social dynamics on Facebook: The Brexit debate. *Social Networks*, 50:6 – 16, 2017.

[53] P. DiMaggio, J. Evans, and B. Bryson. Have American's social attitudes become more polarized? *American Journal of Sociology*, 102(3):690–755, 1996.

[54] R. E. Dunlap, A. M. McCright, and J. H. Yarosh. The political divide on climate change: Partisan polarization widens in the U.S. *Environment: Science and Policy for Sustainable Development*, 58(5):4–23, 2016.

[55] J. Dyer and B. Kolic. Public risk perception and emotion on Twitter during the COVID-19 pandemic. *Applied Network Science*, 5(1):99, 2020.

[56] D. Elgesem, L. Steskal, and N. Diakopoulos. Structure and content of the discourse on climate change in the blogosphere: The big picture. *Environmental Communication*, 9(2):169–188, 2015.

[57] A. M. Ertugrul, Y.-R. Lin, W.-T. Chung, M. Yan, and A. Li. Activism via attention: interpretable spatiotemporal learning to forecast protest activities. *EPJ Data Science*, 8(1):5, 2019.

[58] M. Everett and S. Borgatti. The dual-projection approach for two-mode networks. *Social Networks*, 35(2):204 – 210, 2013. Special Issue on Advances in Two-mode Social Networks.

[59] P. Expert, T. S. Evans, V. D. Blondel, and R. Lambiotte. Uncovering space-independent communities in spatial networks. *Proceedings of the National Academy of Sciences*, 108(19):7663–7668, 2011.

[60] R. Eyre, F. De Luca, and F. Simini. Social media usage reveals recovery of small businesses after natural hazard events. *Nature Communications*, 11(1):1629, 2020.

[61] J. Farrell. Network structure and influence of the climate change countermovement. *Nature Climate Change*, 6(4):370–374, 2016.

[62] J. T. Feezell. Agenda setting through social media: The importance of incidental news exposure and social filtering in the digital era. *Political Research Quarterly*, 71(2):482–494, 2018.

[63] L. Feldman, P. S. Hart, A. Leiserowitz, E. Maibach, and C. Roser-Renouf. Do hostile media perceptions lead to action? The role of hostile media perceptions, political efficacy, and ideology in predicting climate change activism. *Communication Research*, 44(8):1099–1124, 2017.

[64] L. Feldman, E. W. Maibach, C. Roser-Renouf, and A. Leiserowitz. Climate on cable: The nature and impact of global warming coverage on Fox News, CNN, and MSNBC. *The International Journal of Press/Politics*, 17(1):3–31, 2012.

[65] L. Feldman, T. A. Myers, J. D. Hmielowski, and A. Leiserowitz. The mutual reinforcement of media selectivity and effects: Testing the reinforcing spi-

rals framework in the context of global warming. *Journal of Communication*, 64(4):590–611, 2014.

[66] M. Fernandez, L. S. G. Piccolo, D. Maynard, M. Wippoo, C. Meili, and H. Alani. Talking climate change via social media: Communication, engagement and behaviour. In *Proceedings of the 8th ACM Conference on Web Science*, pages 85–94, 2016.

[67] G. Ferraz de Arruda, F. Aparecido Rodrigues, P. Martín Rodríguez, E. Cozzo, and Y. Moreno. A general Markov chain approach for disease and rumour spreading in complex networks. *Journal of Complex Networks*, 6(2):215–242, 2017.

[68] M. A. Fortuna, D. B. Stouffer, J. M. Olesen, P. Jordano, D. Mouillot, B. R. Krasnov, R. Poulin, and J. Bascompte. Nestedness versus modularity in ecological networks: two sides of the same coin? *Journal of Animal Ecology*, 79(4):811–817, 2010.

[69] S. Fortunato and M. Barthélemy. Resolution limit in community detection. *Proceedings of the National Academy of Sciences*, 104(1):36–41, 2007.

[70] D. Freelon, A. Marwick, and D. Kreiss. False equivalencies: Online activism from left to right. *Science*, 369(6508):1197–1201, 2020.

[71] X. Fu, S. Yu, and A. R. Benson. Modelling and analysis of tagging networks in Stack Exchange communities. *Journal of Complex Networks*, 8(5), 2019.

[72] D. Garcia, A. Abisheva, S. Schweighofer, U. Serdült, and F. Schweitzer. Ideological and temporal components of network polarization in online political participatory media. *Policy & Internet*, 7(1):46–79, 2015.

[73] C. Gariazzo, A. Pelliccioni, and M. P. Bogliolo. Spatiotemporal analysis of urban mobility using aggregate mobile phone derived presence and demographic data: A case study in the city of Rome, Italy. *Data*, 4(1):8, 2019.

[74] K. Garimella, G. De Francisci Morales, A. Gionis, and M. Mathioudakis. The effect of collective attention on controversial debates on social media. In *Proceedings of the 2017 ACM on Web Science Conference*, WebSci '17, page 43–52, 2017.

[75] J. P. Gleeson, D. Cellai, J.-P. Onnela, M. A. Porter, and F. Reed-Tsochas. A simple generative model of collective online behavior. *Proceedings of the National Academy of Sciences*, 111(29):10411–10415, 2014.

[76] J. Gomide, A. Veloso, W. Meira, V. Almeida, F. Benevenuto, F. Ferraz, and M. Teixeira. Dengue surveillance based on a computational model of spatio-temporal locality of Twitter. In *Proceedings of the 3rd International Web Science Conference*, WebSci '11, 2011.

[77] N. Grinberg, K. Joseph, L. Friedland, B. Swire-Thompson, and D. Lazer. Fake news on Twitter during the 2016 U.S. presidential election. *Science*, 363(6425):374–378, 2019.

[78] S. Guarino, N. Trino, A. Celestini, A. Chessa, and G. Riotta. Characterizing networks of propaganda on Twitter: a case study. *Applied Network Science*, 5(1):59, 2020.

[79] A. M. Guess, M. Lerner, B. Lyons, J. M. Montgomery, B. Nyhan, J. Reifler, and N. Sircar. A digital media literacy intervention increases discernment between mainstream and false news in the United States and India. *Proceedings of the National Academy of Sciences*, 117(27):15536–15545, 2020.

[80] R. Guimerà, M. Sales-Pardo, and L. A. N. Amaral. Module identification in bipartite and directed networks. *Phys. Rev. E*, 76:036102, 2007.

[81] U. Harush and B. Barzel. Dynamic patterns of information flow in complex networks. *Nature Communications*, 8(1):2181, 2017.

[82] T. Häussler. Heating up the debate? Measuring fragmentation and polarisation in a German climate change hyperlink network. *Social Networks*, 54:303 – 313, 2018.

[83] B. Hawelka, I. Sitko, E. Beinat, S. Sobolevsky, P. Kazakopoulos, and C. Ratti. Geo-located Twitter as proxy for global mobility patterns. *Cartography and Geographic Information Science*, 41(3):260–271, 2014.

[84] X. He and Y.-R. Lin. Measuring and monitoring collective attention during shocking events. *EPJ Data Science*, 6(1):30, 2017.

[85] P. Healy, G. Hunt, S. Kilroy, T. Lynn, J. P. Morrison, and S. Venkatagiri. Evaluation of peak detection algorithms for social media event detection. In *2015 10th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP)*, pages 1–9, 2015.

[86] I. Himelboim, S. McCreery, and M. Smith. Birds of a feather tweet together: Integrating network and content analyses to examine cross-ideology exposure on Twitter. *Journal of Computer-Mediated Communication*, 18(2):40–60, 2013.

[87] J. Hopcroft, O. Khan, B. Kulis, and B. Selman. Tracking evolving communities in large linked networks. *Proceedings of the National Academy of Sciences*, 101(suppl 1):5249–5253, 2004.

[88] M. J. Hornsey, E. A. Harris, and K. S. Fielding. Relationships among conspiratorial beliefs, conservatism and climate scepticism across nations. *Nature Climate Change*, 8(7):614–620, 2018.

[89] Y. Hu, S. Ji, Y. Jin, L. Feng, H. E. Stanley, and S. Havlin. Local structure can identify and quantify influential global spreaders in large scale social networks. *Proceedings of the National Academy of Sciences*, 115(29):7468–7472, 2018.

[90] H. N. Huynh, E. F. Legara, and C. Monterola. A dynamical model of Twitter activity profiles. In *Proceedings of the 26th ACM Conference on Hypertext & Social Media*, HT '15, page 49–57, 2015.

[91] H. Isah, D. Neagu, and P. Trundle. Bipartite network model for inferring hidden ties in crime data. In *Proceedings of the 2015 IEEE/ACM Interna-*

tional Conference on Advances in Social Networks Analysis and Mining 2015, ASONAM '15, pages 994–1001, 2015.

[92] M. Jacomy, T. Venturini, S. Heymann, and M. Bastian. ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PLOS ONE*, 9(6):1–12, 2014.

[93] S. M. Jang and P. S. Hart. Polarized frames on "climate change" and "global warming" across countries and states: Evidence from Twitter big data. *Global Environmental Change*, 32:11–17, 2015.

[94] L. Jasny, J. Waggle, and D. R. Fisher. An empirical examination of echo chambers in US climate policy networks. *Nature Climate Change*, 5:782–786, 2015.

[95] R. Jurdak, K. Zhao, J. Liu, M. AbouJaoude, M. Cameron, and D. Newth. Understanding human mobility from Twitter. *PLOS ONE*, 10(7):1–16, 2015.

[96] E. Katz and P. F. Lazarsfeld. *Personal influence: the part played by people in the flow of mass communications.* Free Press, 1955.

[97] M. J. Keeling and K. T. Eames. Networks and epidemic models. *Journal of The Royal Society Interface*, 2(4):295–307, 2005.

[98] B. W. Kernighan and S. Lin. An efficient heuristic procedure for partitioning graphs. *The Bell System Technical Journal*, 49(2):291–307, 1970.

[99] A. P. Kirilenko and S. O. Stepchenkova. Public microblogging on climate change: One year of Twitter worldwide. *Global Environmental Change*, 26:171 – 182, 2014.

[100] M. Kitsak, L. K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H. E. Stanley, and H. A. Makse. Identification of influential spreaders in complex networks. *Nature Physics*, 6(11):888–893, 2010.

[101] V. Koh, W. Li, G. Livan, and L. Capra. Offline biases in online platforms: a study of diversity and homophily in Airbnb. *EPJ Data Science*, 8(1):11, 2019.

[102] J. Kunegis, A. Lommatzsch, and C. Bauckhage. The Slashdot zoo: Mining a social network with negative edges. In *Proceedings of the 18th International Conference on World Wide Web*, WWW '09, page 741–750, 2009.

[103] V. Lampos and N. Cristianini. Tracking the flu pandemic by monitoring the social web. In *2010 2nd International Workshop on Cognitive Information Processing*, pages 411–416, 2010.

[104] V. Larivière, Y. Gingras, and É. Archambault. The decline in the concentration of citations, 1900–2007. *Journal of the American Society for Information Science and Technology*, 60(4):858–862, 2009.

[105] D. B. Larremore, A. Clauset, and A. Z. Jacobs. Efficiently inferring community structure in bipartite networks. *Phys. Rev. E*, 90:012805, 2014.

[106] M. Latapy. Main-memory triangle computations for very large (sparse (power-law)) graphs. *Theoretical Computer Science*, 407(1):458 – 473, 2008.

[107] M. Latapy, C. Magnien, and N. D. Vecchio. Basic notions for the analysis of large two-mode networks. *Social Networks*, 30(1):31–48, 2008.

[108] D. Lazer, A. Pentland, L. Adamic, S. Aral, A.-L. Barabási, D. Brewer, N. Christakis, N. Contractor, J. Fowler, M. Gutmann, T. Jebara, G. King, M. Macy, D. Roy, and M. Van Alstyne. Computational social science. *Science*, 323(5915):721–723, 2009.

[109] K. Lee, J. Caverlee, K. Y. Kamath, and Z. Cheng. Detecting collective attention spam. In *Proceedings of the 2nd Joint WICOW/AIRWeb Workshop on Web Quality*, WebQuality '12, page 48–55, 2012.

[110] S. Lee, S.-i. Song, M. Kahng, D. Lee, and S.-g. Lee. Random walk based entity ranking on graph for multidimensional recommendation. In *Proceedings of the Fifth ACM Conference on Recommender Systems*, pages 93–100, 2011.

[111] J. Lehmann, B. Gonçalves, J. J. Ramasco, and C. Cattuto. Dynamical classes of collective attention in Twitter. In *Proceedings of the 21st International Conference on World Wide Web*, WWW '12, page 251–260, 2012.

[112] A. Leiserowitz, E. Maibach, C. Roser-Renouf, G. Feinberg, and P. Howe. Global Warming's Six Americas: September 2012. `http://climatecommunication.yale.edu/publications/global-warmings-six-americas-in-september-2012/`, 2013. Accessed: 2020-01-10.

[113] Y. Li and C. You. What is the difference of research collaboration network under different projections: Topological measurement and analysis. *Physica A: Statistical Mechanics and its Applications*, 392(15):3248 – 3259, 2013.

[114] Y.-R. Lin, B. Keegan, D. Margolin, and D. Lazer. Rising tides or rising stars?: Dynamics of shared attention on Twitter during media events. *PLOS ONE*, 9(5):1–12, 2014.

[115] Y.-R. Lin, D. Margolin, B. Keegan, A. Baronchelli, and D. Lazer. #Bigbirds never die: Understanding social dynamics of emergent hashtags. *Seventh International AAAI Conference on Weblogs and Social Media*, 2013.

[116] X. Long, L. Jin, and J. Joshi. Exploring trajectory-driven local geographic topics in Foursquare. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, UbiComp '12, page 927–934, 2012.

[117] P. Lorenz-Spreen, B. M. Mønsted, P. Hövel, and S. Lehmann. Accelerating dynamics of collective attention. *Nature Communications*, 10(1):1759, 2019.

[118] P. Lorenz-Spreen, F. Wolf, J. Braun, G. Ghoshal, N. Djurdjevac Conrad, and P. Hövel. Tracking online topics over time: understanding dynamic hashtag communities. *Computational Social Networks*, 5(1):9, 2018.

[119] X. Lu and B. K. Szymanski. A regularized stochastic block model for the robust community detection in complex networks. *Scientific Reports*, 9(1):13247, 2019.

[120] Z. Lu, J. Wahlström, and A. Nehorai. Community detection in complex networks via clique conductance. *Scientific Reports*, 8(1):5982, 2018.

[121] J. Maddock, K. Starbird, H. J. Al-Hassani, D. E. Sandoval, M. Orand, and R. M. Mason. Characterizing online rumoring behavior using multidimensional signatures. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, CSCW '15, page 228–241, 2015.

[122] F. M. D. Marquitti, P. R. Guimarães, M. M. Pires, and L. F. Bittencourt. MODULAR: software for the autonomous computation of modularity in large network sets. *Ecography*, 37(3):221–224, 2014.

[123] A. N. Medvedev, J.-C. Delvenne, and R. Lambiotte. Modelling structure and predicting dynamics of discussion threads in online boards. *Journal of Complex Networks*, 7(1):67–82, 2018.

[124] D. Melamed. Community structures in bipartite networks: A dual-projection approach. *PLOS ONE*, 9(5):1–5, 2014.

[125] D. Melamed, A. Harrell, and B. Simpson. Cooperation, clustering, and assortative mixing in dynamic networks. *Proceedings of the National Academy of Sciences*, 115(5):951–956, 2018.

[126] M. Mendoza, B. Poblete, and I. Valderrama. Nowcasting earthquake damages with Twitter. *EPJ Data Science*, 8(1):3, 2019.

[127] S. Messing and S. J. Westwood. Selective exposure in the age of social media: Endorsements trump partisan source affiliation when selecting news online. *Communication Research*, 41(8):1042–1063, 2014.

[128] J. Metag, T. Füchslin, and M. S. Schäfer. Global warming's five Germanys: A typology of Germans' views on climate change and patterns of media use and information. *Public Understanding of Science*, 26(4):434–451, 2017.

[129] S. Mirzaee and Q. Wang. Urban mobility and resilience: exploring Boston's urban mobility network through Twitter data. *Applied Network Science*, 5(1):75, 2020.

[130] T. Mitra, G. Wright, and E. Gilbert. Credibility and the dynamics of collective attention. *Proc. ACM Hum.-Comput. Interact.*, 1(CSCW), 2017.

[131] J. Möller, D. Trilling, N. Helberger, and B. van Es. Do not blame it on the algorithm: an empirical assessment of multiple recommender systems and their impact on content diversity. *Information, Communication & Society*, 21(7):959–977, 2018.

[132] I. Moutidis and H. T. P. Williams. Utilizing complex networks for event detection in heterogeneous high-volume news streams. In *Complex Networks and Their Applications VIII*, pages 659–672, 2020.

[133] S. A. Myers, A. Sharma, P. Gupta, and J. Lin. Information network or social network? The structure of the Twitter follow graph. In *Proceedings of the 23rd International Conference on World Wide Web*, WWW '14 Companion, page 493–498, 2014.

[134] M. Newman. Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, 46(5):323–351, 2005.

[135] M. E. J. Newman. Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Phys. Rev. E*, 64:016132, 2001.

[136] M. E. J. Newman. Assortative mixing in networks. *Phys. Rev. Lett.*, 89:208701, 2002.

[137] M. E. J. Newman. Fast algorithm for detecting community structure in networks. *Phys. Rev. E*, 69:066133, 2004.

[138] M. E. J. Newman. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23):8577–8582, 2006.

[139] M. E. J. Newman. Equivalence between modularity optimization and maximum likelihood methods for community detection. *Phys. Rev. E*, 94:052315, 2016.

[140] M. E. J. Newman and M. Girvan. Finding and evaluating community structure in networks. *Phys. Rev. E*, 69, 2004.

[141] M. E. J. Newman, S. H. Strogatz, and D. J. Watts. Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E*, 64:026118, 2001.

[142] T. P. Newman. Tracking the release of IPCC AR5 on Twitter: Users, comments, and sources following the release of the Working Group I summary for policymakers. *Public Understanding of Science*, 26(7):815–825, 2017.

[143] A. Olteanu, C. Castillo, N. Diakopoulos, and K. Aberer. Comparing events coverage in online news and social media: The case of climate change. In *International AAAI Conference on Web and Social Media*, 2015.

[144] S. O'Neill, H. T. P. Williams, T. Kurz, B. Wiersma, and M. Boykoff. Dominant frames in legacy and social media coverage of the IPCC Fifth Assessment Report. *Nature Climate Change*, 5, 2015.

[145] S. J. O'Neill and M. Hulme. An iconic approach for representing climate change. *Global Environmental Change*, 19(4):402 – 410, 2009.

[146] L. Page, S. Brin, R. Motwani, and T. Winograd. The PageRank citation ranking: Bringing order to the web. Technical Report 1999-66, Stanford InfoLab, November 1999.

[147] C. M. Peak, A. Wesolowski, E. zu Erbach-Schoenberg, A. J. Tatem, E. Wetter, X. Lu, D. Power, E. Weidman-Grunewald, S. Ramos, S. Moritz, C. O. Buckee, and L. Bengtsson. Population mobility reductions associated with travel restrictions during the Ebola epidemic in Sierra Leone: use of mobile phone data. *International Journal of Epidemiology*, 47(5):1562–1570, 2018.

[148] T. P. Peixoto. Hierarchical block structures and high-resolution model selection in large networks. *Phys. Rev. X*, 4:011047, 2014.

[149] T. P. Peixoto. Disentangling homophily, community structure and triadic closure in networks. `https://arxiv.org/abs/2101.02510`, 2021.

[150] C. Poletto, S. V. Scarpino, and E. M. Volz. Applications of predictive modelling early in the COVID-19 epidemic. *The Lancet Digital Health*, 2(10):e498–e499, 2020.

[151] A. J. Porter and I. Hellsten. Investigating participatory dynamics through social media using a multideterminant "frame" approach: The case of Climategate on YouTube. *Journal of Computer-Mediated Communication*, 19(4):1024–1041, 2014.

[152] M. Porter. An algorithm for suffix stripping. *Program: electronic library and information systems*, 14(3):130–137, 1980.

[153] L. Prandi and G. Primiero. Effects of misinformation diffusion during a pandemic. *Applied Network Science*, 5(1):82, 2020.

[154] H. A. Prasetya and T. Murata. A model of opinion and propagation structure polarization in social media. *Computational Social Networks*, 7(1):2, 2020.

[155] W. Rand, J. Herrmann, B. Schein, and N. Vodopivec. An agent-based model of urgent diffusion in social media. *Journal of Artificial Societies and Social Simulation*, 18(2):1, 2015.

[156] S. Romano, N. X. Vinh, J. Bailey, and K. Verspoor. Adjusting for chance clustering comparison measures. *Journal of Machine Learning Research*, 17(134):1–32, 2016.

[157] D. M. Romero, B. Meeder, and J. Kleinberg. 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*, WWW '11, page 695–704, 2011.

[158] S. Saito, Y. Hirata, K. Sasahara, and H. Suzuki. Tracking time evolution of collective attention clusters in Twitter: Time evolving nonnegative matrix factorisation. *PLOS ONE*, 10(9):1–17, 2015.

[159] F. Saracco, M. J. Straka, R. D. Clemente, A. Gabrielli, G. Caldarelli, and T. Squartini. Inferring monopartite projections of bipartite networks: an entropy-based approach. *New Journal of Physics*, 19(5):053022, 2017.

[160] K. Sasahara. Visualizing collective attention using association networks. *New Generation Computing*, 34(4):323–340, 2016.

[161] K. Sasahara, Y. Hirata, M. Toyoda, M. Kitsuregawa, and K. Aihara. Quantifying collective attention from tweet stream. *PLOS ONE*, 8(4):1–10, 2013.

[162] D. A. Scheufele and N. M. Krause. Science audiences, misinformation, and fake news. *Proceedings of the National Academy of Sciences*, 116(16):7662–7669, 2019.

[163] F. Schlosser, B. F. Maier, O. Jack, D. Hinrichs, A. Zachariae, and D. Brockmann. COVID-19 lockdown induces disease-mitigating structural changes in mobility networks. *Proceedings of the National Academy of Sciences*, 117(52):32883–32890, 2020.

[164] A. L. Schmidt, F. Zollo, M. Del Vicario, A. Bessi, A. Scala, G. Caldarelli, H. E. Stanley, and W. Quattrociocchi. Anatomy of news consumption on Facebook. *Proceedings of the National Academy of Sciences*, 114(12):3035–3039, 2017.

[165] J. P. Schuldt, S. H. Konrath, and N. Schwarz. "Global warming" or "climate change"? Whether the planet is warming depends on question wording. *Public Opinion Quarterly*, 75(1):115–124, 2011.

[166] A. Segerberg and W. L. Bennett. Social media and the organization of collective action: Using Twitter to explore the ecologies of two climate change protests. *The Communication Review*, 14(3):197–215, 2011.

[167] E. Shearer and J. Gottfried. *News Use Across Social Media Platforms 2017*. Pew Research Center, 2017. `http://www.journalism.org/2017/09/07/news-use-across-social-media-platforms-2017/`.

[168] E. Shearer and K. E. Matsa. *News Use Across Social Media Platforms 2018*. Pew Research Center, 2018. `https://www.journalism.org/2018/09/10/news-use-across-social-media-platforms-2018/`.

[169] J. Shin and K. Thorson. Partisan selective sharing: The biased diffusion of fact-checking messages on social media. *Journal of Communication*, 67(2):233–255, 2017.

[170] M. Spruce, R. Arthur, and H. T. P. Williams. Using social media to measure impacts of named storm events in the United Kingdom and Ireland. *Meteorological Applications*, 27(1):e1887, 2020.

[171] A. Srivastava, A. J. Soto, and E. Milios. Text clustering using one-mode projection of document-word bipartite graphs. In *Proceedings of the 28th Annual ACM Symposium on Applied Computing*, SAC '13, pages 927–932, 2013.

[172] K. Starbird. Examining the alternative media ecosystem through the production of alternative narratives of mass shooting events on Twitter. In *Proc. International AAAI Conference on Web and Social Media*, 2017.

[173] K. Starbird and L. Palen. (How) will the revolution be retweeted? Information diffusion and the 2011 Egyptian uprising. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, CSCW '12, page 7–16, 2012.

[174] C. R. Sunstein. *Republic.com 2.0*. Princeton University Press, 2007.

[175] M. Tambuscio, G. Ruffo, A. Flammini, and F. Menczer. Fact-checking effect on viral hoaxes: A model of misinformation spread in social networks. In *Proceedings of the 24th International Conference on World Wide Web*, WWW '15 Companion, page 977–982, 2015.

[176] L. Taylor. No place to hide? The ethics and analytics of tracking mobility using mobile phone data. *Environment and Planning D: Society and Space*, 34(2):319–336, 2016.

[177] M. ten Thij, A. Kaltenbrunner, D. Laniado, and Y. Volkovich. Collective attention patterns under controlled conditions. *Online Social Networks and Media*, 13:100047, 2019.

[178] J. H. Tien, M. C. Eisenberg, S. T. Cherng, and M. A. Porter. Online reactions to the 2017 'Unite the right' rally in Charlottesville: measuring polarization in Twitter networks using media followership. *Applied Network Science*, 5(1):10, 2020.

[179] P. Valkenburg and J. Peter. Comm Research—Views from Europe— Five Challenges for the Future of Media-Effects Research. *International Journal of Communication*, 7(0), 2013.

[180] L. van der Maaten and G. Hinton. Visualizing high-dimensional data using t-SNE. *Journal of Machine Learning Research*, 9:2579–2605, 2008.

[181] D. Vasques Filho and D. R. J. O'Neale. Degree distributions of bipartite networks and their projections. *Phys. Rev. E*, 98:022307, 2018.

[182] S. Vosoughi, D. Roy, and S. Aral. The spread of true and false news online. *Science*, 359(6380):1146–1151, 2018.

[183] L.-Z. Wang, Z.-D. Zhao, J. Jiang, B.-H. Guo, X. Wang, Z.-G. Huang, and Y.-C. Lai. A model for meme popularity growth in social networking systems based on biological principle and human interest dynamics. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(2):023136, 2019.

[184] Z. Wang, T. Hou, D. Song, Z. Li, and T. Kong. Detecting Review Spammer Groups via Bipartite Graph Projection. *The Computer Journal*, 59(6):861–874, 2016.

[185] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442, 1998.

[186] I. S. Weaver. Preferential attachment in randomly grown networks. *Physica A: Statistical Mechanics and its Applications*, 439:85 – 92, 2015.

[187] I. S. Weaver, H. Williams, I. Cioroianu, L. Jasny, T. Coan, and S. Banducci. Communities of online news exposure during the UK General Election 2015. *Online Social Networks and Media*, 10-11:18 – 30, 2019.

[188] I. S. Weaver, H. Williams, I. Cioroianu, M. Williams, T. Coan, and S. Banducci. Dynamic social media affiliations among UK politicians. *Social Networks*, 54:132–144, 2018.

[189] I. S. Weaver, H. T. P. Williams, and R. Arthur. A social Beaufort scale to detect high winds using language in social media posts. *Scientific Reports*, 11(1):3647, 2021.

[190] B. E. Weeks, D. S. Lane, D. H. Kim, S. S. Lee, and N. Kwak. Incidental exposure, selective exposure, and political information sharing: Integrating online exposure patterns and expression on social media. *Journal of Computer-Mediated Communication*, 22(6):363–379, 2017.

[191] J. A. Weill, M. Stigler, O. Deschenes, and M. R. Springborn. Social distancing responses to COVID-19 emergency declarations strongly differentiated by income. *Proceedings of the National Academy of Sciences*, 117(33):19658–19660, 2020.

[192] L. Weng, A. Flammini, A. Vespignani, and F. Menczer. Competition among memes in a world with limited attention. *Scientific Reports*, 2(1):335, 2012.

[193] L. Weng, F. Menczer, and Y.-Y. Ahn. Virality prediction and community structure in social networks. *Scientific Reports*, 3(1):2522, 2013.

[194] L. Weng, F. Menczer, and Y.-Y. Ahn. Predicting successful memes using network and community structure. In *International AAAI Conference on Web and Social Media*, 2014.

[195] A. Wesolowski, C. O. Buckee, K. Engø-Monsen, and C. J. E. Metcalf. Connecting Mobility to Infectious Diseases: The Promise and Limits of Mobile Phone Data. *The Journal of Infectious Diseases*, 214(suppl_4):S414–S420, 2016.

[196] H. T. Williams, J. R. McMurray, T. Kurz, and F. H. Lambert. Network analysis reveals open forums and echo chambers in social media discussions of climate change. *Global Environmental Change*, 32:126 – 138, 2015.

[197] M. Williams, I. Cioroianu, and H. Williams. Different news for different views: Political news-sharing communities on social media through the UK General Election in 2015. In *Proc. Workshop on News and Public Opinion (NECO) at International AAAI Conference on Web and Social Media*, 2016.

[198] S. Wojcik and A. Hughes. *Sizing Up Twitter Users.* Pew Research Center, 2019. `https://www.pewresearch.org/internet/2019/04/24/sizing-up-twitter-users/`.

[199] F. Wu and B. A. Huberman. Novelty and collective attention. *Proceedings of the National Academy of Sciences*, 104(45):17599–17601, 2007.

[200] J. Wyse, N. Friel, and P. Latouche. Inferring structure in bipartite networks using the latent blockmodel and exact ICL. *Network Science*, 5(1):45–69, 2017.

[201] E. Yan and Y. Ding. Scholarly network similarities: How bibliographic coupling networks, citation networks, cocitation networks, topical networks, coauthorship networks, and coword networks relate to each other. *Journal of the American Society for Information Science and Technology*, 63(7):1313–1326, 2012.

[202] L. Yang, T. Sun, M. Zhang, and Q. Mei. We know what you #tag: Does the dual role affect hashtag adoption? In *Proceedings of the 21st International Conference on World Wide Web*, WWW '12, page 261–270, 2012.

[203] L. E. Young, E. Sidnam-Mauch, M. Twyman, L. Wang, J. J. Xu, M. Sargent, T. W. Valente, E. Ferrara, J. Fulk, and P. Monge. Disrupting the COVID-19 misinfodemic with network interventions: Network solutions for network problems. *American Journal of Public Health*, 111(3):514–519, 2021.

[204] S. D. Young, C. Rivers, and B. Lewis. Methods of using real-time social media technologies for detection and remote monitoring of HIV outcomes. *Preventive Medicine*, 63:112–115, 2014.

[205] B. Zhang, S. van der Linden, M. Mildenberger, J. R. Marlon, P. D. Howe, and A. Leiserowitz. Experimental effects of climate messages vary geographically. *Nature Climate Change*, 8(5):370–374, 2018.

[206] X. Zhang, C. Moore, and M. E. J. Newman. Random graph models for dynamic networks. *The European Physical Journal B*, 90(10):200, 2017.

[207] Y. Zhang. Language in our time: An empirical analysis of hashtags. In *The World Wide Web Conference*, WWW '19, page 2378–2389, 2019.

[208] C. Zhao, A. Zeng, and C. H. Yeung. Characteristics of human mobility patterns revealed by high-frequency cell-phone position data. *EPJ Data Science*, 10(1):5, 2021.

[209] Z. Zhao, J. Zhao, Y. Sano, O. Levy, H. Takayasu, M. Takayasu, D. Li, J. Wu, and S. Havlin. Fake news propagates differently from real news even at early stages of spreading. *EPJ Data Science*, 9(1):7, 2020.

[210] C. Zhou, L. Feng, and Q. Zhao. A novel community detection method in bipartite networks. *Physica A: Statistical Mechanics and its Applications*, 492:1679–1693, 2018.

[211] M. Zitnik, R. Sosič, and J. Leskovec. Prioritizing network communities. *Nature Communications*, 9(1):2544, 2018.