

Who Says What with Whom

Using Bi-Spectral Clustering to Organize and Analyze Social Media Protest Networks

Kenneth Joseph, Ryan J. Gallagher, Brooke Foucault Welles

CCR 2 (2): 153–174

DOI: 10.5117/CCR2020.2.002,JOSE

Abstract

Social media protest networks involve many participants, from long-time activists to individuals who are engaged only in a particular protest event. We propose a new approach to studying how various communities of users participate in protest events. Our approach combines two methodological innovations. First, rather than study tweets central to one event, we collect full timelines of user activity leading up to participation in a focal event. Second, we propose bi-spectral clustering as a scalable computational method for rapidly identifying communities of users by the words (hashtags) they use. Using a large sample of tweets from users who discussed the 2016 protests in Charlotte, North Carolina following the extrajudicial killing of Keith Lamont Scott as a case study, we demonstrate how bi-spectral clustering can be applied to sort, sample, and identify ideologically and thematically coherent clusters whose members participated in the protest on Twitter. Our proposed approach provides another tool in the mixed methods scholar's toolkit to computationally sort and cluster large-scale network data by allowing researchers to look beyond focal hashtags or keywords and situate protest messages within the broader communication context of participating users.

Keywords: social media and protest, social networks, Twitter, protest networks, spectral methods

A growing body of research suggests that social media, especially Twitter, has become a key resource for advancing social change (Castells, 2015;

Gerbaudo, 2012; Jackson and Foucault Welles, 2015; Papacharissi & de Fatima Oliveira, 2012; Tufekci, 2017). Much existing work in this space has focused on discussion around individual hashtags (e.g. #BlackLivesMatter) and/or select users (e.g. activist accounts) to examine how online activism and activists work. Although such analyses have produced valuable insights, they are somewhat limited in their abilities to speak about broader context and temporal dynamics. More specifically, we know that hashtags connect conversations across multiple events (Freelon, McIlwain, & Clark, 2016b), and users come and go from conversations around hashtags (Budak & Watts, 2015; Conover et al., 2013; Freelon, McIlwain, & Clark, 2016a). Further, the real power of online activism is – arguably – not the ability to trend a particular hashtag, but the ability for everyday people to sustain an idea and draw connections between incidents over time (Jackson, Bailey, and Foucault Welles, 2020).

Yet, these dynamics are infrequently tracked. We posit that this limitation is not a theoretical one, but rather a methodological one. It is challenging to track a set of users and their conversations leading up to a focal event. In this paper, we introduce an approach to data collection from Twitter, that expands the scope of content typically used to analyze online activism. Specifically, we outline a process to identify a set of users relevant to a focal event and then use the Twitter API to collect the full set of these users' tweets, rather than those tweets relevant only to the focal event. In doing so, we lessen, to some degree, the extent to which we are “selecting on the dependent variable” (Tufekci, 2014) during our analyses; capturing instead a bigger picture of how engagement with a particular activism hashtag fits into the broader range of a Twitter user's content production. Moreover, by leveraging full timelines of tweets we can situate users into communities, and examine how those communities engaged with the event.

A critical question, however, is how to define these communities. Prior work has generally relied heavily on computational network analysis (Conover et al., 2011; Freelon et al., 2016b; González-Bailón & Wang, 2016), identifying communities of users through structural patterns in their social interactions. In the present work, we instead rely on the content of users' tweets to identify communities. More specifically, we rely on the fact that users tweet hashtags as a means of signifying actual or desired membership in social and/or topical communities on Twitter (Yang, Sun, Zhang, & Mei, 2012). Because of this, we can identify communities by the hashtags they share, and define the ideological stance of that community by its representative hashtags. Underlying this use of hashtags as community markers is the longstanding sociological theory of the duality of people and culture

(Breiger & Puetz, 2015). Duality theory argues that people are defined by cultural artifacts they produce, and simultaneously that cultural artifacts are defined by the people who leverage them. Consequently, duality theory implies that a community can be identified and understood via structural properties of the network created by people (users) leveraging distinct cultural artifacts (hashtags). In other words, we posit that clustering people by hashtags is a meaningful way to understand communities on Twitter, as communities can be meaningfully defined by the things they say.

We use a method from the document clustering literature – bi-spectral clustering (Dhillon, 2001) – to identify these communities of users and the hashtags that best define them. Bi-spectral clustering provides a scalable and relatively simple method for identifying sub-communities based on who is in them and what those individuals are talking about. More generally, it serves as a means of exploring large-scale datasets and uncovering subsets of users who engage with a topic of interest in different ways, as well as those that do not engage with the topic at all (irrelevant users) and those that engage with it speciously (e.g. for marketing, misinformation, etc., as is the case with many bot accounts).

In the sections that follow, we describe bi-spectral clustering in detail and provide an overview of related methods. Then, we apply the technique to a sample of tweets about the November, 2016 protests in Charlotte, North Carolina, after Keith Lamont Scott was killed by a Charlotte police officer. We show how bi-spectral clustering can be used to filter data to a subset of particular interest, to “drill down” hierarchically on this data to further understand interesting ideological inconsistencies in the results, and to observe small, cohesive subsets of users that might be overlooked in broader studies of activism on large datasets. We conclude with a discussion of the broader implications of bi-spectral clustering, including ethical considerations and other applications where bi-spectral clustering may be useful. Code used to replicate our approach with other datasets can be found in the public release accompanying this work.¹

Bi-spectral Clustering

Method Description

Bi-spectral clustering was developed by Dhillon (2001) to cluster text documents based on the words used within each document. The technique relies on the construction of *bipartite graphs*. In network science, a *graph* consists of a set of objects (people, documents, organizations, etc.) connected

by the relationships between them (friendship, common words, collaborators, etc.). In a bipartite graph, there is more than one type of object – for example, people connected to social groups (e.g. Borgatti & Halgin, 2011), or editors connected to Wikipedia articles (e.g. Keegan, Gergle, & Contractor, 2011). Relationships in bipartite graphs exist only between objects of different types. Although Twitter data are often arranged into single-object networks (e.g. user-user, hashtag-hashtag) they can easily be defined as bipartite graphs (e.g. user-hashtag) as well. Here, then, users are the “documents” and hashtags are the “words.” (Benigni et al., 2017).

The novelty of Dhillon’s method is that it not only clusters text documents that use similar words together, it clusters the documents together *with the words that best define these documents*. Thus, one could run the method on a collection of social protest articles and find one cluster of mass communication articles with the words “framing” and “mainstream news”, while another cluster might contain sociology articles grouped together with the terms “class” and “race”. In the context of our use case of user-hashtag networks, we can simultaneously uncover collections of users with the hashtags that best characterize their Twitter communication.

Formally, bi-spectral clustering seeks to find a co-clustering of a bipartite graph (here, users and hashtags). First, we define a bipartite graph where the nodes are users ($(u_1, u_2, \dots, u_n \in U)$) and hashtags ($(h_1, h_2, \dots, h_m \in H)$). An edge in the graph, E_{ij} , is formed when a user u_i expresses a hashtag h_j in one or more of her tweets; the weight of E_{ij} is the number of times u_i uses h_j in her tweets.²

When formulated like this, a natural way to think about the optimal clustering of users and hashtags is to find the clustering that minimizes the weight of edges across clusters, ensuring we form clusters with a high density of internal edge weights and few connections outside the cluster. More formally, if we wish to uncover a set of K clusters of users and hashtags, we would like to minimize the sum of the *cut-weights* across all pairs of clusters. The cut (weight) between two clusters C_1 and C_2 , composed of both users and hashtags, is defined as $(C_1, C_2) = \sum_{u_i \in C_1, h_j \in C_2} E_{ij}$. To minimize the cut weight of the cuts across all K clusters, we therefore would like to minimize the quantity $(C_1, C_2, C_3, \dots, C_K) = \sum_{x < y} \text{cut}(C_x, C_y)$.

The naive solution to the problem of minimizing cut scores across a set of K clusters is almost always simply to have $K - 1$ clusters with a single object (user or hashtag) in each, and all other objects in the one other cluster. While this is mathematically near optimal in many cases, this is undesirable for substantive social questions, because it would not allow us to study meaningful aggregations of the data. To avoid this solution, Dhillon

(2001), therefore, suggests weighting each cluster to ensure that the method is likely to generate a reasonable number of data points in each cluster we specify. Mathematically, the ideal clustering under this paradigm should minimize the quantity $cut_{weighted}(C_1, C_2, C_3, \dots, C_K) = \sum_{x < y} cut_{weighted}(C_x, C_y)$, where $cut_{weighted}(C_1, C_2)$ is defined as

$$cut_{weighted}(C_1, C_2) = \frac{cut(C_1, C_2)}{\sum_{y \in C_1} weight(y)} + \frac{cut(C_1, C_2)}{\sum_{z \in C_2} weight(z)}$$

and $weight(y) = \sum_k E_{yk}$ if y is a user and $\sum_j E_{jy}$ if y is a hashtag:

Having established the function to minimize in order to find a meaningful clustering of the bipartite graph of interest, we might now hope to simply produce a clustering of users and hashtags by maximizing the above function. Unfortunately, however, doing so is computationally intractable.³ Dhillon's innovation was to show that one could find a suitable, efficient and theoretically appealing relaxation to this clustering problem (equivalently, to this maximization problem) using straightforward matrix algebra and the well-known k -means algorithm (Hartigan & Wong, 1979). We refer readers to the original text for details on the mathematics of this clustering.

Related Approaches

Bi-spectral clustering is directly related to three lines of work. First, the underlying mathematical model is similar to many standard approaches in the social sciences. For example, the mathematics underlying bispectral clustering are similar to correspondence analysis (CA); however, CA assumes a categorical, rather than a continuous, distribution of edge weights in the bipartite graph. Second, the formulation of the problem as a bipartite graph clustering implies a relation to community detection algorithms in the network science literature. While most community detection algorithms are not designed to account for the two-mode nature of bipartite networks recent work has cast the problem of clustering documents and words as a question of finding a maximally compact encoding of the document-word bipartite network (Gerlach, Peixoto, & Altmann, 2018). While this approach is intriguing and warrants future inspection within the context of computational communication research, here we focus on bi-spectral clustering, which leverages a lighter, more readily accessible mathematical framework to determine textual clusters. The third line of work is the host of methods that have been developed since Dhillon's work to computationally analyze text and documents. We discuss here three of the

most popular approaches in the literature – latent Dirichlet allocation (LDA; Blei et al., 2003), the structural topic model (STM; Roberts et al., 2014), and the word2vec word embedding model (Mikolov et al., 2013) – and how they relate to and diverge from bi-spectral clustering.

The mathematical model used by bi-spectral clustering is similar to the mathematics underlying the creation of word embeddings. Both are based largely on singular value decomposition, explicitly in the bi-spectral clustering case and either implicitly or explicitly in the case of word embeddings (Levy and Goldberg, 2014). However, while popular for other tasks, word embedding models like word2vec (Mikolov et al., 2013) are not conducive to the clustering of users and hashtags that we are interested in here. Word embeddings focus on learning latent representations of words (hashtags) only, rather than jointly learning representations, let alone clusters, of users and hashtags. Further, they do so by analyzing the word by word (hashtag by hashtag) matrix, rather than the user by hashtag matrix. To use word embeddings for the purpose of clustering users and hashtags, then, one would have to carry out an intermittent step of creating representations of users by, e.g., averaging the hashtags they use. This is essentially an ad-hoc version of bi-spectral clustering without any of the theoretical validation provided by Dhillon (2001), plus an ill-defined intermediate step.

Both LDA and the STM are Bayesian ad-mixture models, better known in the text analysis literature as topic models. Topic models have become a tool of choice for both computational scientists and social scientists keen on interpreting themes in large corpora of text (e.g. DiMaggio, 2015; DiMaggio, Nag, & Blei, 2013). Topic models share several similarities with bi-spectral clustering – both can be used to identify clusters of users and hashtags in a bipartite network. A critical difference, however, is that in a topic model, one assumes that users are “mixtures” of communities, and that communities are “mixtures” of hashtags. In other words, topic models estimate the probability that each hashtag and each user is associated with each “topic” (community). Each user, therefore, might be a member of multiple communities, each of which is defined by a set of hashtags, which itself might represent a variety of communities.

Consider, for example, a large network of people discussing contemporary political issues. One person in that network may be a Black Lives Matter activist, a transgender woman, and a survivor of sexual violence. Using a topic model, that user could potentially be characterized as a mixture of three communities (topics) where #BlackLivesMatter, #GirlsLikeUs, and #MeToo are top hashtags that characterize each community respectively. However, if another user in the same broad network can be

characterized by two communities, one conservative (#tcot) and one “color-blind” (#AllLivesMatter), the mathematics underlying topic models will force the first user to be represented as a mixture of these communities as well. As a result, every user is represented as a mixture of every community in the network to some extent, which reduces the probabilistic weight that can be given to the communities that best characterize any particular user.

In an online appendix provided in the code release for this paper, we show how this probabilistic assumption, characteristic of both LDA and the STM, can introduce difficulties in analysis for mixed methods researchers. In the online appendix, we apply latent Dirichlet allocation (LDA) to the same dataset we use in our case study, showing how this results in the need for a series of ad hoc decisions not required from bi-spectral clustering that reduce the coherence and interpretability of user and hashtag clusters. We do not consider a comparison to the STM because it faces the same concerns, despite differing in two major ways from LDA. First, the STM allows for incorporation of external covariates (e.g. time, political affiliation) that can vary with topics. While useful in other contexts, these covariates are a part of the same probabilistic framework of LDA. Second, the STM builds off the correlated topic model (CTM; Blei and Lafferty, 2007), and uses a prior that allows for correlations between topics that are assigned to a document. In contrast, LDA assumes topics are independently assigned. Again though, adding topic correlations does not change the ad-mixture nature of the STM. The shared assumptions of LDA, the STM, and the CTM are more critical in the context of the present work than their differences. More specifically, the difficulties we identify with LDA for the analyses of interest necessarily exist for the CTM and STM as well. For this reason, our online appendix only looks at LDA and the qualitative difficulties that arise when using it to co-cluster users and hashtags.

The primary difference between topic models and bispectral clustering, then, is the assumption that users and hashtags belong to multiple groups (LDA and the STM) versus users and hashtags belonging to a single primary group (bispectral clustering). The assumption that users and hashtags are assigned to only a single community can be problematic, as we know individuals exist within multiple social circles (Szell, Lambiotte, & Thurner, 2010). Leveraging this assumption therefore requires that we think in a new way about individuals spanning multiple groups as interstitial users. Instead of individuals spanning multiple communities, bi-spectral clustering enforces the assumption that individuals have a primary community, which they may extend beyond at various times. Leaning on duality theory, we argue that it is appropriate to assign people to “home” clusters based on

the hashtags they typically use over a long period of time, while allowing for the possibility that they may occasionally temporarily participate in other clusters by using their hashtags for various reasons. As we will show in the case study, analyzing the data with this lens can provide clues as to where users attempted to span community boundaries for activist purposes.

Sample Application: Charlotte Protests

In the late afternoon of September 20, 2016, Keith Lamont Scott, a 43-year old Black resident of Charlotte, North Carolina was shot and killed by a Charlotte-Mecklenburg police officer as he exited his car. Hours later, citizens of Charlotte gathered in the streets to protest Scott's death. Echoing themes from a series of protests about the extrajudicial killing of Black men by police officers, Charlotte protesters (online and in the streets) connected with the larger Black Lives Matter movement, including protests over Terence Crutcher's fatal shooting by a police officer in Tulsa, Oklahoma and Tyre King's fatal shooting by a Columbus, Ohio police officer just days before. The protests in Charlotte continued throughout the nights of September 20 and 21 and concluded when Charlotte Mayor Jennifer Roberts imposed a citywide curfew in the overnight hours of September 22-23.

To illustrate how bi-spectral clustering works to identify communities participating in a large online protest, we assembled a set of 43,514 Twitter users who tweeted using keywords related to the Charlotte protests during the time of the protests (September 20-22, 2016) as well as up to the last 3,200 tweets they sent⁴ prior to the protest. Data collection proceeded in three phases. First, we collected tweets that were potentially relevant to the protests. We then collected the timelines of all users who participated in the focal event. Finally, we used two strategies to filter the data to exclude noisy data (e.g. from bots) and individuals who were not part of the conversation.

Collecting focal event tweets

Our dataset of Charlotte protest tweets was drawn from two sources. First, starting on September 21, 2016 and ending two weeks later, we collected data from the Streaming API using the search terms *charlotte*, *#Charlotte*, *#KeithLamontScott*, *#CLT*, *#CharlotteProtest*, *#KeithScott*, *#CharlotteRiot*, *#CharlottePD*, and *#CharlotteUprising* using Twitter's publicly available tools to do so.⁵ In addition to this real-time collection, we leveraged a historical archive of Twitter data available to certain institutions that provides a

pseudo-random 10% sample of all tweets sent on Twitter. From this dataset, we extracted any tweets from September 20-22, 2016 that included protest-related search terms (Table 2) and added the tweets to the data collected from the Streaming API. In total, the combination of these two approaches to obtaining tweets relevant to the protests resulted in a set of 3,257,253 tweets from 143,915 users. These users are the participants in the focal (protest) event.

Identifying focal users and collecting their timelines

To associate them with communities, we assembled tweet timelines for each focal event participant. First, using the Twitter Search API, we pulled down up to the last 3,200 tweets posted by each of the 143,916 users in our dataset. For a small subset of very active Twitter users, this set of 3,200 tweets did not stretch back to a period before the time of the protests. For these accounts, such as the account for activist Deray McKesson, we supplemented the set of tweets we had for these users from the Advanced Search page on Twitter,⁶ which allows one to see all non-retweets from a given user (beyond this 3,200-tweet limit) from September 1, 2016 and onward. Doing so allowed us to perform our clustering on not only the focal Charlotte tweets, but also on all, or nearly all, tweets generated by users who tweeted during the Charlotte protests. This allowed us to cluster those users in terms of their broader ideas and practices (as expressed by their other tweets before the protests) and also as discussants of the protest itself.

Data Cleaning

We took several additional steps to subsample a collection of users that were engaged in discussions around the focal event. First, we lower-cased all hashtags in the data. This was the only preprocessing step performed on the text. Second, we used bi-spectral clustering to identify users who used relevant keywords but were unrelated to the focal protest. For instance, we found clusters of users discussing Princess Charlotte (daughter of Prince William and Kate, Duke and Duchess of Cambridge), an animated show entitled *Charlotte*, and a professional wrestler named Charlotte. We also saw clusters of bot accounts that used the hashtags we searched as a way to promote irrelevant content (e.g. for selling natural supplements). We removed users from these clusters from our analyses (from 143,915 users to 91,828 users, a reduction of almost 40% of users with off-topic or spam content), helping to focus results. Note that the use of bi-spectral clustering was not critical in this step – another clustering approach or bot-removal techniques could have been used. However, bispectral clustering did allow

us to exclude from our analysis clusters of both users and hashtags that, as a whole, were not of interest for further study.

Finally, as a filter on the set of users that were actively engaged in Charlotte protest discussions, we further restricted our analysis to the set of users who sent at least one tweet during the protest that was retweeted at least one time. Although this strategy may have excluded some users who made relevant contributions, it also engages the users themselves as gatekeepers of relevant content; a strategy that is common in determining the boundaries of online protest networks (e.g. Jackson and Foucault Welles, 2015; 2016). Setting the filter at the user (rather than individual tweet) level ensures that a diversity of message types remain in the corpus for analysis. As nearly all tweets are retweeted within the first 24 hours after they are sent (Kwak, Lee, Park, & Moon, 2010; Starbird & Palen, 2012), we determined whether or not a tweet was retweeted by checking retweet counts for all tweets in our 3.2M tweet sample one week after data collection ended. Of the 91,828 users produced after our initial spam and off-topic removal step, 47,388 (51.6%) had at least one tweet that had been retweeted. Of the 47,388 accounts selected for further analysis, 43,514 used at least one hashtag in any of their tweets. We additionally restricted our analysis to a set of 137,401 hashtags that were used by at least ten users during this time period, as hashtags used by fewer numbers of users are unlikely to be interesting for analysis. Thus, our final bipartite graph to cluster is a graph with 43,514 users and 137,401 hashtags. On average, users expressed 1,275 total hashtags (median 719) and 368 unique hashtags (median 260).

Bi-Spectral Clustering Analysis

Our analysis is separated into two subsections. We first provide a brief overview of the application of bi-spectral clustering to the dataset described above. Second, we provide a more in-depth discussion of several clusters identified by bi-spectral clustering, focusing on the types of questions that bi-spectral clustering can help answer when examining focal protest events within the context of user timelines.

Overview of Bi-spectral clustering of Charlotte Protest Data

We ran bi-spectral clustering on the cleaned data with $k=100$, producing 100 clusters of users and hashtags. We decided on $k=100$ after carrying out two analyses. The first was a qualitative assessment of the most popular hashtags in each cluster at $k=10$, 25, 50, and 200. We found that 100 clusters

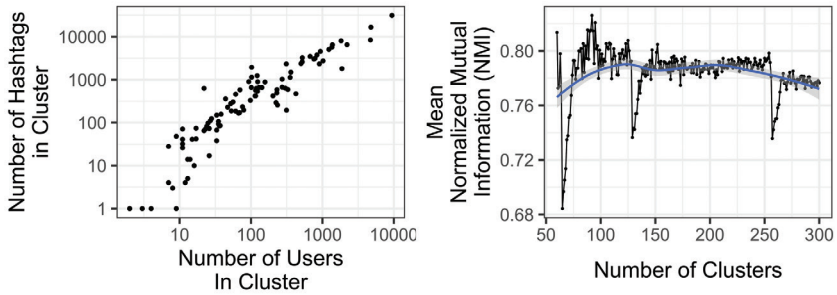


Figure 1. Left: Users and hashtags per cluster for 100 clusters produced using bi-spectral clustering on Charlotte protest tweets. Each point is a cluster, where the horizontal axis is the number of users in the cluster, and vertical axis is the number of hashtags. Right: The average Normalized Mutual Information (NMI, y-axis) between a given number of clusters k (x-axis) and the previous clusterings $k-1, k-2, \dots, k-10$.

provided more nuanced and potentially meaningful separations of the data than smaller numbers of clusters, but that at 200 the clusters had too few users to meaningfully study. The second was a quantitative analysis of the Normalized Mutual Information (McDaid, Greene, & Hurley, 2001) for different numbers of clusters. Specifically, we first compute clusterings for $k=50$ to $k=300$. For each cluster size k , we then compute the average Normalized Mutual Information, a measure of clustering consistency, for the clustering obtained with k clusters and $k-1, k-2, \dots, k-10$ clusterings. Doing so gives a measure of how consistent clustering is across different values of k . Results are presented in the right-hand plot in Figure 1. The figure shows that clusterings are generally consistent with previous clusterings, except when a threshold is crossed where an additional dimension is added to the SVD (recall that we use $\text{ceiling}(\log_2(k))$ dimensions). Further, we find that the highest levels of consistency peak between values of k between 75-125, suggesting that 100 is a reasonable number of clusters to select.

The left-hand plot in Figure 1 shows that the clusters varied widely in size, a typical result for bi-spectral clustering. While the majority of clusters had between 50-1000 users (median 73, mean 435) and 100-1000 hashtags (median 297, mean 1495), this distribution was skewed – for example, the top 5 clusters in terms of overall size contain 52.7% of all users.

As an initial analysis step, we manually inspected the top 25 most-used hashtags for each cluster. Even in the cleaned data, the clusters extracted from the data exposed users that were still engaged in discussions not relevant to the protest. These clusters of users and hashtags were filtered out.

For instance, on the left in Table 2, we see representative hashtags for a cluster of 273 users who are focused on professional wrestling. Because these users are artifacts induced by the polysemy of the keywords we used to search Twitter, they can easily be removed from the analysis. In the right-hand column of Table 2, we see representative hashtags for a set of marketing bots. Because bi-spectral clustering allows us to see what users said before and during the protest events, these bots emerge clearly as off-topic from the protest itself (even if they did hijack relevant trending hashtags to promote their goods and services). Again, such users are highly unlikely to be of interest, and thus can be safely removed.

Table 2. Top five hashtags from clusters eliminated from further analysis of Charlotte protest tweets

Professional Wrestling Cluster	Bot Cluster
<i>wrestlemania</i>	<i>bakersfield</i>
<i>wwebattleground</i>	<i>las</i>
<i>wwedraft</i>	<i>erie</i>
<i>mitb</i>	<i>apartment</i>
<i>royalrumble</i>	<i>lansing</i>

This filtering process illustrates a benefit of our approach – we were able to rapidly identify (and eliminate) tweets that were not relevant to our subsequent analysis. Researchers often spend a great deal of time cleaning data in cases where relevant hashtags or keywords are common words, as in this Charlotte case, or where hashtags were used for multiple purposes (e.g. #GirlsLikeUs was used by Black trans women and also to promote a never-produced Taylor Swift movie). Further, bots that are difficult to detect when only considering a focal hashtag often appear in their own clusters under this method, which considers their historical communication. Used iteratively, bi-spectral clustering allows us to rapidly find collections of users that could pollute an in-depth analysis of online activism. In this sense, the method can simply be used as a filtering step before further analyses, as we do here and in our initial data cleaning. However, the method also produces cohesive clusters of users and hashtags that are of theoretical interest (described below).

Qualitative Analysis of #Charlotte Protest Participation

We now turn to how the clusters identified by bi-spectral clustering can facilitate further in-depth qualitative analysis, illustrating with examples from the Charlotte protest tweets. As with any qualitative analysis informed by computational techniques, our goal is not a census of Charlotte protest discourse, but

rather an illustration of three key strengths of bi-spectral clustering within a mixed-methods framework (in addition to the sorting and filtering strengths discussed above). In-depth qualitative research always depends on working with small samples. Mixed-methods studies of online protest are often informed by a combination of computational methods to identify interesting cases and in-depth qualitative analysis to interpret the substantive importance of those cases (e.g. Freelon, Lopez, Clark & Jackson, 2018; Jackson, Bailey, & Foucault Welles, 2020). The following three examples highlight ways that qualitative analysis of users and hashtags can be uniquely informed by sorting protest data with bi-spectral clustering of users and hashtags.

Many different communities of users participated in the Charlotte protests, including the mainstream media, users providing tactical updates about protest activities, users tweeting about systemic racism, users tweeting in support of the protests in Charlotte and elsewhere around the country, political tweets aligned with liberal and conservative ideologies, Black Lives Matter activists and supporters, clusters of users tweeting about or on behalf of institutions located in Charlotte, and more. Because bi-spectral clustering creates communities based on an entire history of tweets, we are uniquely able to see how people who *typically* discuss particular topics *specifically* engage in Charlotte protest tweets. In what follows we examine three clusters in more detail to illustrate how this typical-specific dichotomy that bi-spectral clustering provides can effectively inform nuanced critical and qualitative analyses.⁷

Black Lives Matter Activists and Charlotte Newcomers: Distinguishing Core Activists and Peripheral Supporters

As one in a series of protests about police violence against Black citizens over the course of several years, it is unsurprising that a cluster of committed Black Lives Matter activists produced Charlotte protest messages. Table 3 shows the top five most-used hashtags during the protest (out of 3,478 in the cluster) and top five most-followed users (out of 672 users) of a large cluster of Black Lives Matter activists who participated in the Charlotte protest network.

Table 3. Top hashtags and users in the Black Lives Matter Activist cluster

Hashtags	Users
<i>tyreking</i>	<i>deray</i>
<i>charlotteuprising</i>	<i>ShaunKing</i>
<i>nojusticenopeace</i>	<i>Blklivesmatter</i>
<i>shutitdown</i>	<i>OccupyWallStNYC</i>
<i>stl</i>	<i>BreeNewsome</i>

Experts in online activism will note that, the cluster includes activists such as Deray McKesson, Shaun King, and Bree Newsome who are known for tweeting during protest events. We see from their hashtags that they tweeted about the protests in Charlotte, and also connected those protests to a broader pattern protest and police violence against Black people elsewhere around the country. For instance, on the morning of September 21 activist Deray McKesson tweeted “Have all those good cops people keep talking about released a public statement about #TerenceCrutcher or #KeithLamontScott yet?” highlighting the similarities in the circumstances of both murders while also pushing back against a common defense used to minimize the scope of police violence (e.g. the problem is “just a few bad cops” who kill Black people.) This discursive strategy has emerged as typical of long-term online racial justice activists who aim to connect each new murder to a pattern of racialized state violence (Jackson, Bailey & Foucault Welles, 2020).

This cluster was distinct from a related cluster of users whose first protest tweets appeared during the Charlotte protests. This cluster included more of the “everyday citizens” found to play central roles in hashtag networks analyzed using different methods (e.g. Jackson and Foucault Welles, 2016). For example, on September 22, connecting the deaths of several Black men killed by police officers in September of 2016, a Black woman tweeted,⁸ “I have not even processed the deaths of #KeithLamontScott & #TerrenceCrutcher and now there’s already another one. RIP #TawonBoyd.” Although this and other messages in the cluster were semantically and discursively similar to those in the Black Lives Matter activist cluster, these users showed less history of tweeting about protests prior to the events in Charlotte. This result underscores the distinct advantage of coupling our approach to data collection of user timelines, rather than just event tweets, with the bi-spectral clustering method. Namely, we are able to distinguish thematically similar tweets into distinct clusters, based on users’ historical tendencies to use similar language. In doing so, we can distinguish the core of longer-term Black Lives Matter activists from newcomers engaged in online racial justice activism for the first time in response to Keith Lamont Scott’s murder. Research suggests both types of users play important – but different – roles in spreading protest messages (Barberá et al., 2015) and bi-spectral clustering uniquely allows us to easily differentiate the two groups.

News Cluster: Multiple Resolutions through Iterative Clustering

Table 4 presents the top five users and hashtags from another cluster, the largest produced by our method. Broadly consisting of journalists and news

organizations covering breaking news, including news of the protests in Charlotte, the *News* cluster contains 9,406 users and 31,259 hashtags.

Table 4. Top hashtags and users in the News cluster

<i>Hashtags</i>	<i>Users</i>
<i>blacklivesmatter</i>	<i>nytimes</i>
<i>debatenight</i>	<i>cnn</i>
<i>trump</i>	<i>TheEconomist</i>
<i>charlotte</i>	<i>funnyordie</i>
<i>breaking</i>	<i>TIME</i>

The ease and speed of bi-spectral clustering allows us to further refine the cluster by simply re-running the method on this subsample of users and hashtags. To perform this “drill down”, we sub-select from our data only those users and hashtags in this cluster and run bi-spectral cluster as above, except allowing for only 25 clusters. Results in Table 5 show evidence of a subset of users more closely aligned with the Black Lives Matter messaging evidenced in the activist cluster discussed above. These users include progressive media outlets, Black journalists and celebrities, and some left-leaning mainstream news outlets.

Table 5. Top five hashtags and users in the progressive cluster within the news community

<i>Hashtags</i>	<i>Users</i>
<i>blacklivesmatter</i>	<i>williecolon</i>
<i>charlotte</i>	<i>TheAtlantic</i>
<i>breaking</i>	<i>thenation</i>
<i>keithlamontscott</i>	<i>hollyrpeete</i>
<i>charlotteprotest</i>	<i>xonecole</i>

Although thematically similar to the activist cluster, these progressive news users can be differentiated from users in the activist cluster in two ways. First, users in this subset of the news cluster tended to use more mainstream terms to describe protest events (e.g. *#blacklivesmatter* and *#charlotte*), while those in the activist cluster used more terms associated with nonviolent resistance (e.g. *#shutitdown* and *#nojusticenopeace*). Second, unlike the activist cluster which made connections between Charlotte and other protests around the country, this subset of news cluster more often made connections with the broader political context during the time of the protests (e.g. *#donaldtrump* and *#imwithher*, references to the 2016

U.S. presidential campaigns). This suggests different orientations to power and social change and stronger projection between the news cluster users and mainstream power structures than we saw within the activist cluster. Indeed, the inclusion of mainstream news outlets and reporters, along with Black celebrities like actress Holly Robinson Peete, who regularly engages politicians in discussion of racial justice and other issues, further supports these alignments.

These results underscore both the methodological utility of bi-spectral clustering and an interesting avenue for future work. With respect to the former, results suggest that the method is useful for taking a hierarchical perspective on ideological diversity around a particular context. Here, we showed how we can “drill down” within a collection of users defined by their use of popular newsworthy hashtags to those that also align with activist content. In this way, we can find people who use similar language to describe protest events but who (potentially) have different ideological takes on how to address the issues at hand. While there are variants of LDA that could produce this hierarchy algorithmically (Griffiths, Jordan, Tenenbaum, & Blei, 2004), they come at the expense of heavier mathematical and computational machinery without resolving the fundamental probabilistic concerns addressed earlier.

Further, the results of the iterative bi-spectral clustering suggest important future ways to analyze possible bridges between thematically similar but discursively distinct sub-clusters of communication networks, including this case of activists and mainstream press. By iteratively applying bi-spectral clustering, we can find users saying similar things to different audiences, potentially identifying people and messages that can bridge otherwise disparate communities over time.

Local Institutions: Brokerage Across Clusters

The final cluster we consider is a set of 2,338 hashtags and 906 users that suggests how combining bi-spectral clustering results with more in-depth qualitative analysis may lead to interesting new paths for analyses. Table 6 displays the top five hashtags and users from this cluster. We see that the hashtags characterizing this group focus on institutions relevant to the larger Charlotte area, including universities (#unc) local trade organizations (#ncga), and, the Carolina Panthers, an NFL football team located in Charlotte. Similarly, the top users in the cluster include local news organizations and personalities, as well as the official Twitter account for the Carolina Panthers.

Table 6. Top five users and hashtags in the local Charlotte cluster

<i>Hashtags</i>	<i>Users</i>
<i>panthers</i>	<i>Panthers</i>
<i>keepounding</i>	<i>WRAL</i>
<i>clt</i>	<i>wsocvtv</i>
<i>ncpol</i>	<i>theobserver</i>
<i>unc</i>	<i>newsobserver</i>

This cluster presents an interesting subset of users who are not commonly included in discussions of protest events, nor are they particularly active in protest discussions, relative to other clusters such as those discussed above. Instead, they are the institutions that serve the local community that was the geographic site of the protests. We see evidence of their sustained commitment to community mission, for instance when the local library tweeted information about services available during the protest and curfew. We also see clear examples of brokerage across clusters; for instance, people using the terms most relevant to Charlotte institutions to advance activist causes. Recall that, based on her twitter timeline, activist Bree Newsome is a member of the Black Lives Matter activist cluster. But, on September 24, in advance of a Carolina Panthers football game and following several days of protests in Charlotte, Newsome tweeted, “Will the.@Panthers #KneelForCharlotte? That’s a powerful way players can show support for the communities hurting this week. #KeepPounding.” Referencing the actions of professional football player Colin Kaepernick who took a knee (rather than standing) during the National Anthem at several 2016 pre-season games to protest police violence against Black people, Newsome used the language of the local institutions cluster (Panthers hashtag #KeepPounding) to reach a new audience. While the tweet is explicitly an invitation for the local professional football team and players to engage with the local community and protest events, it also implicitly invites the broader community of those interested in the football team (but perhaps unaware of the protests, which had largely subsided before game day) to reflect on the protest events and the challenges facing the local community.

This example illustrates an interesting point for further study that would not otherwise be visible in an analysis that focused on tweets sent during the protest alone. Had we formed clusters based on discourse during the protests alone, Newsome’s football-related discourse would have sorted her into the “local institutions” cluster, which is correct if we only focus on the focal event. However, she is widely known, and active on Twitter, as an activist. Clustering her based on her broader discourse allows us to

do so appropriately and observe how she transcended clusters to effect social change. More generally, the ability for bi-spectral clustering to produce communities based on shared communication history allows researchers to study how usually-distinct communities intersect during protests, including specific cases of brokerage like the one illustrated here.

Discussion and Future Work

Bi-spectral clustering, combined with the collection of data from the entirety of users' timelines, offers a new way to collect and analyze large Twitter and other social media text datasets for cohesive ideological clusters of users and hashtags. Considering full user timelines allows us to place users in the context of broader discursive communities as they interact during a particular focal event. Bi-spectral clustering then can be used to home in on particular subcollections of users and/or hashtags for further quantitative and qualitative analyses, and also filter out unwanted content. The method is easy to implement, open-source, and scalable, allowing for rapid, principled, iterative analyses of large datasets in a mixed methods framework. Together, accounting for the discursive histories of users through their timelines and co-clustering users and hashtags according to those histories provide a new angle for understanding how focal protest events unfold.

Of course, the proposed methodology is not without its limitations. First, while we attempted to limit the bias in our dataset due to sampling the Twitter Streaming API with a set of keywords, this process is likely to have missed a number of Twitter users who were expressing their opinions related to the Charlotte protests via other terminology. As methods exist to update keyword searches in real-time (Linder, 2017), future work could utilize them to construct a better sample. Second, while bi-spectral clustering is a principled algorithm for reliably defining and jointly clustering users and the hashtags they use, we have not focused on showing with certainty it is *better* than other methods, instead showing that it bypasses certain ad hoc thresholding decisions of applying LDA in this context. Finally, as with any method that uncovers otherwise-hidden connections in social media texts, researchers using bi-spectral clustering must consider how to ethically handle issues of privacy and consent to participate in research. We have taken extra care in this paper not to reveal the specific content of tweets generated by individuals who are not in the public eye. Although this is not the only precaution we could have taken, we felt this maintained an appropriate balance of transparency in our scientific process, while still

maintaining the privacy of those especially unlikely to be aware of the possible implications of their tweets. We encourage researchers to carefully consider the ethical implications of reporting on the results of bi-spectral clustering before applying the method.

Despite these limitations and challenges, bi-spectral clustering is a useful technique for organizing, sampling, and analyzing large-scale social media text data. Although our case study focuses on activism networks, for which we believe the method is especially useful, bi-spectral clustering could be applied to any large dataset that includes users generating text about topics of interest. As illustrated in our examples, it may be especially useful for distinguishing new users from those who have a more sustained history of communicating about a particular topic, drilling down to sub-communities of related but discursively distinct users, and identifying cases where users intentionally bridge between communities to advance new messages. As a tool for quickly processing and grouping user-generated text, bi-spectral clustering is a useful addition to the computational communication researcher's toolkit.

Notes

- 1 https://github.com/kennyjoseph/bispectral_comm
- 2 One could potentially add in edges between users that, e.g., mention each other, or between hashtags that, e.g., co-occur in the same tweets. However, prior work has suggested that doing so actually has a negative impact on performance; therefore, we here focus only on the bipartite representation (Benigni, Joseph, & Carley, 2017)
- 3 That is, it is an NP-complete problem. See, e.g., Krishnamurthy (1985) for an introduction to computational complexity and NP-completeness
- 4 And more in some cases, as described below
- 5 <https://github.com/twitter/hbc>
- 6 <https://twitter.com/search-advanced?lang=en>
- 7 Made available as part of the Github release for this article at the link in Footnote 1
- 8 To protect the privacy of this woman, who is not a celebrity or public figure, we have elected not to include her Twitter handle here. Her identity was confirmed through a visual inspection of her account and a second social media account linked in her Twitter profile. The text of her tweet has been altered to prevent reverse searching without altering its meaning; see Ayers et al. (2018) for further context on this decision.

References

- Arora, S., Ge, R., Halpern, Y., Mimno, D., Moitra, A., Sontag, D., ... Zhu, M. (2012). *A Practical Algorithm for Topic Modeling with Provable Guarantees* (ArXiv E-Print No. 1212.4777).

- Barberá, P., Wang, N., Bonneau, R., Jost, J. T., Nagler, J., Tucker, J., & González-Bailón, S. (2015). The Critical Periphery in the Growth of Social Protests. *PLOS ONE*, *10*(11), e0143611.
- Benigni, M. C., Joseph, K., & Carley, K. M. (2017). Online extremism and the communities that sustain it: Detecting the ISIS supporting community on Twitter. *PLOS ONE*, *12*(12), e0181405.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research.*, *3*, 993-1022.
- Blei, D. M., & Lafferty, J. D. (2007). A correlated topic model of science. *The Annals of Applied Statistics*, *1*(1), 17-35.
- Borgatti, S. P., & Halgin, D. S. (2011). Analyzing affiliation networks. *The Sage Handbook of Social Network Analysis*, *1*, 417-433.
- Breiger, R. L., & Puetz, K. (2015). Culture and networks. *The International Encyclopedia of the Social and Behavioral Sciences*, 557-62.
- Budak, C., & Watts, D. J. (2015). Dissecting the spirit of Gezi: Influence vs. selection in the Occupy Gezi movement. *Sociological Science*, *2*, 370-397.
- Castells, M. (2015). *Networks of outrage and hope: Social movements in the Internet age*. John Wiley & Sons.
- Conover, M. D., Davis, C., Ferrara, E., McKelvey, K., Menczer, F., & Flammini, A. (2013). The Geospatial Characteristics of a Social Movement Communication Network. *PLOS ONE*, *8*(3), e55957.
- Conover, M., Ratkiewicz, J., Francisco, M. R., Gonçalves, B., Menczer, F., & Flammini, A. (2011). Political polarization on twitter. *Proceedings of the International Conference on the Web and Social Media*, *133*, 89-96.
- Dhillon, I. S. (2001). Co-clustering Documents and Words Using Bipartite Spectral Graph Partitioning. *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 269-274.
- DiMaggio, P. (2015). Adapting computational text analysis to social science (and vice versa). *Big Data & Society*, *2*(2), 2053951715602908.
- DiMaggio, P., Nag, M., & Blei, D. (2013). Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of U.S. government arts funding. *Poetics*, *41*(6), 570-606.
- Freelon, D., Lopez, L., Clark, M. D., & Jackson, S. J. (2018). *How Black Twitter and other social media communities interact with mainstream news*. Miami, FL, USA: Knight Foundation.
- Freelon, D., McIlwain, C., & Clark, M. (2016a). Quantifying the power and consequences of social media protest. *New Media & Society*, 1461444816676646.
- Freelon, D., McIlwain, C. D., & Clark, M. (2016b). Beyond the hashtags: #Ferguson, #BlackLivesMatter, and the online struggle for offline justice. *Center for Media & Social Impact, American University*.
- Gerbaudo, P. (2012). *Tweets and the Streets*. Pluto Press London.
- Gerlach, M., Peixoto, T. P., & Altmann, E. G. (2018). A network approach to topic models. *Science Advances*, *4*(7), eaaq1360.
- González-Bailón, S., & Wang, N. (2016). Networked discontent: The anatomy of protest campaigns in social media. *Social Networks*, *44*, 95-104.
- Griffiths, T. L., Jordan, M. I., Tenenbaum, J. B., & Blei, D. M. (2004). Hierarchical topic models and the nested Chinese restaurant process. *Advances in Neural Information Processing Systems*, 17-24.
- Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, *28*(1), 100-108.

- Jackson, S.J., Bailey, M. & Foucault Welles, B. (2020). *#HashtagActivism: Networks of Race and Gender Justice*. Cambridge, MA: MIT Press.
- Jackson, S. J., & Foucault Welles, B. (2015). Hijacking# myNYPD: Social media dissent and networked counterpublics. *Journal of Communication*, 65(6), 932-952.
- Jackson, S. J., & Foucault Welles, B. (2016). # Ferguson is everywhere: Initiators in emerging counterpublic networks. *Information, Communication & Society*, 19(3), 397-418.
- Keegan, B., Gergle, D., & Contractor, N. (2011). Hot off the wiki: dynamics, practices, and structures in Wikipedia's coverage of the Tōhoku catastrophes. *Proceedings of the 7th International Symposium on Wikis and Open Collaboration*, 105-113. ACM.
- Kwak, H., Lee, C., Park, H., & Moon, S. (2010). What is Twitter, a social network or a news media? *Proceedings of the 19th International Conference on World Wide Web*, 591-600.
- Levy, O., & Goldberg, Y. (2014). Neural word embedding as implicit matrix factorization. In *Advances in neural information processing systems* (pp. 2177-2185).
- Linder, F. (2017). *Improved Data Collection from Online Sources Using Query Expansion and Active Learning*.
- McDaid, A. F., Greene, D., & Hurley, N. (2011). Normalized mutual information to evaluate overlapping community finding algorithms. *arXiv preprint arXiv:110.2515*.
- Morstatter, F., & Liu, H. (2017). In search of coherence and consensus: measuring the interpretability of statistical topics. *The Journal of Machine Learning Research*, 18(1), 6177-6208.
- Morstatter, F., Pfeffer, J., Liu, H., & Carley, K. M. (2013). Is the sample good enough? comparing data from twitter's streaming api with twitter's firehose. *Proceedings of the International Conference on the Web and Social Media*.
- Papacharissi, Z., & de Fatima Oliveira, M. (2012). Affective News and Networked Publics: The Rhythms of News Storytelling on #Egypt. *Journal of Communication*, 62(2), 266-282.
- Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., ... Rand, D. G. (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science*, 58(4), 1064-1082.
- Schmidt, B. M. (2012). Words alone: Dismantling topic models in the humanities. *Journal of Digital Humanities*, 2(1), 49-65.
- Starbird, K., & Palen, L. (2012). (How) will the revolution be retweeted?: information diffusion and the 2011 Egyptian uprising. *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, 7-16.
- Szell, M., Lambiotte, R., & Thurner, S. (2010). Multirelational organization of large-scale social networks in an online world. *Proceedings of the National Academy of Sciences*, 107(31), 13636-13641.
- Tufekci, Z. (2014). Big Questions for Social Media Big Data: Representativeness, Validity and Other Methodological Pitfalls. *ICWSM '14: Proceedings of the 8th International AAAI Conference on Weblogs and Social Media*.
- Tufekci, Z. (2017). *Twitter and tear gas: The power and fragility of networked protest*. Yale University Press.
- Wallach, H., Mimno, D., & McCallum, A. (2009). Rethinking LDA: Why priors matter. *Advances in Neural Information Processing Systems*, 22, 1973-1981.
- Yang, L., Sun, T., Zhang, M., & Mei, Q. (2012). We know what @you #tag: does the dual role affect hashtag adoption? *Proceedings of the 21st International Conference on World Wide Web*, 261-270.

About the authors

Kenneth Joseph, University at Buffalo, Computer Science and Engineering Department, Buffalo, NY.

Correspondence address: University at Buffalo, The State University of New York, 338 Davis Hall, Box 602000, Buffalo, NY 14260-2500 (kjoseph@buffalo.edu)

Ryan J. Gallagher, Northeastern University, Network Science Institute, Boston, MA.

Brooke Foucault Welles, Northeastern University, Network Science Institute and Department of Communication, Boston, MA.