A Network Analysis of Black Lives Matter Protests Using Twitter Data

Jennifer Allen

Yale University
Abstract

This paper does a network analysis of the Black Lives Matters movement using data collected from four Twitter protests taking place during March and April of 2016. The analysis performed was two-fold. First, a multivariable regression model was developed to determine which variables best predicted how influential a user was in recruiting other users to the protest. Second, a case study analysis of the Andrew Loku protest was done in order to reveal more information about these influential players and to study how protest activity unfolded. Results showed that being mentioned frequently by other users, tweeting more about the Black Lives Matters movement, and having a high eigenvector centrality score predicted higher levels of influence. However, having more followers did not predict higher levels of influence in this network. Additionally, the case study analysis showed that these influential users were a mix of official Black Lives Matters accounts, media accounts, and regular user accounts, reflecting the ways in which Black Lives Matters is both a grassroots movement and an organized protest network.
Introduction

Since their inception in the late 2000s, social media platforms such as Facebook and Twitter have emerged as revolutionary tools that promote modern social activism. Indeed, in 2011, The Washington Post coined the term “Hashtag Activism” to describe the #OccupyWallStreet protest and other internet-based social movements that make use of Twitter hashtags. Since then, the term has been applied to a litany other social movements that have galvanized the online community, including #BringBackOurGirls, #Kony2012, and #YesAllWomen. Twitter has emerged as more than just an outlet for millennials to broadcast their thoughts, but as an organizing platform for grassroots activists.

This rise of new platforms like Twitter has challenged assumptions from older literature on collective action and protest recruitment. Online protest activity on Twitter differs in many ways from other forms of social activism. First, the creation of an online platform eliminates many of the costs of protest participation. Whereas participants in the Freedom Summer risked their lives while attempting to register black voters, Twitter activists face far less severe consequences. The barriers to participation in online movements are much lower than in the “real world.” Second, Twitter lacks the organizational structure typically found in examples of successful protest recruitment. Twitter is a decentralized network and does not appear to have established communities or meso-level organizations that prior research shows play an important role in protest recruitment (McAdam, 1986). Yet, despite this apparent limitation, Twitter has proved to be an enormously fruitful breeding ground for protest activity.

Because of its popularity and novelty, Twitter has been a vibrant area of study for sociologists studying social networks. Many of these studies have focused on how information disseminates on Twitter and which users are influential in sparking informational cascades
(Bakshy et al, 2011; Bakshy et al, 2012; Cheng et al, 2014; Hong & Davidson, 2010). Bakshy et al (2011) found that users who have been influential in the past and users with high numbers of followers are most successful at triggering large cascades of information flow. Additionally, several other studies have looked at the particular role that Twitter played in protest movements such as the Arab Spring, the Indignados movement in Spain, and the Occupy Wall Street movement (Bastos, Mercea & Charpentier, 2015; Gonzalez-Bailon, Borge-Holthoefer, Moreno, 2011; Russell, 2011; Wilson & Dunn, 2011). Researchers have found that collective action on Twitter defies expectations of classical theory of collective action that assumes that rational actors perform cost/benefit analysis to decide to participate (Bastos, Mercea & Charpentier, 2015; Gonzalez-Bailon, Borge-Holthoefer, Moreno, 2011). Instead, online protest activity bears more resemblance to Granovetter’s threshold model of collective action. Movement participation of Twitter users is better described in terms of social influence and as a function of how many of their peers have already participated (Granovetter, 1978; Gonzalez-Bailon, Borge-Holthoefer, Moreno, 2011).

Most of the literature has focused on movements at their inception, and little research has been done about protest movements after their initial recruitment phase. Additionally, while many studies have looked at the network dynamics of Twitter protests, few studies have identified which factors make some users better at recruiting participants than others. In this paper, I focus on how Twitter operates as an organizing network for a mature movement and which users are most influential within that network.

The Black Lives Matter movement began in 2012 in response to the death of the black teen Trayvon Martin by neighborhood watchman George Zimmerman. In the wake of the event, hashtags such as #justicefortrayvon and #blacklivesmatter began trending on Twitter. What
began as decentralized Twitter activity then transformed into international protest movement advocating against police brutality and violence against black people. The movement evolved to be more than just a reaction to a specific incident and has grown into both an online and offline network of like-minded activists.

Indeed, as the movement has matured, it has begun to formally organize. The movement now has an official website which defines itself as a “chapter-based national organization working for the validity of Black life.” However, despite the official website’s claims about the movement’s overarching structure, the effectiveness of that structure has been questioned. An article published in The Daily Beast entitled “Who Really Runs #BlackLivesMatter?” details the controversy. The article points out the ambiguity as to what constitutes being part of the “official” Black Lives Matters movement and how different participants in the movement often spread contradictory messages (Collins & Mak, 2015). The founder of the official movement states that only chapters who have registered with the national organization are part of the official movement while simultaneously claiming, “anyone can be a Black Lives Matter activist.” There are many questions and conflicting claims about who is running the movement and how the movement’s message is spread.

**Research Question and Hypothesis**

Focusing on Black Lives Matter movement on Twitter opens many interesting questions for research. First of all, which users are influential within the Black Lives Matters network? What are their characteristics? Are they regular users, or are they part of the official “Black Lives Matters” organization? How does protest information spread on this network? This paper will attempt to address some of these questions by analyzing Twitter data related to recent activity of the Black Lives Matters movement and the follower network derived from that data.
The first question I seek to answer is which types of users are most influential at sparking protest activity within the Black Lives Matters movement. Prior research has shown that users with many followers and users who are referenced frequently by other users in the network (through either @mentions or retweets) are typically most successful at sparking information cascades. Thus, I hypothesize that users with high numbers of followers and high numbers of mentions will be most successful at influencing other users to join protest activity.

Also, since the Black Lives Matters movement is an older movement than previously studied in the literature, I hypothesize that users who had previously been more active in the larger Black Lives Matters movement will be more successful at influencing other users to join new protest activity. I use the number of past tweets that a user has posted using the #blacklivesmatters or #blm hashtags as a measure of their activity in the movement.

In addition, I will identify who the most influential nodes in the network are. I want to determine whether the most influential nodes are official Black Lives Matters accounts (for example @BLM_TO, the official account of the Black Lives Matters movement in Toronto), self-identified Black Lives Matters activists, or users without connection to the greater movement.

Finally, I will analyze the dynamics of this network. I hypothesize that online protest activity is largely driven by central actors with low thresholds to participate, where a threshold is defined as the number of people who the user follows who join the protest movement before the user. I hypothesize that the threshold distribution for users looks similar to other online movements.

Data and Methodology
I collected data of Twitter activity relating to four Black Lives Matters protests taking place in March and April of 2016. Protests were selected from https://elephrame.com/textbook/BLM, a site that keeps a continuously updated list of all Black Lives Matters protests. I define joining the protest as the user sending his or her first tweet related to the protest.

For the first stage of data collection and construction, I identified several hashtags related to each protest. I then used the Twitter API to collect tweets that contain those hashtags and key terms. Unfortunately, Twitter’s rate limits only allow collection of a week’s worth of historical data on a particular hashtag. However, the hashtags on Twitter that were examined did not typically trend for periods longer than a week, since I was looking at smaller protests related to the larger Black Lives Matter movement. In total, 4316 of tweets were collected from 2417 users relating to four different protests.

The second stage of data collection and construction consisted of recreating the follower network for each of these protests. The follower graph was constructed by adding directed edges between each user’s followers and friends (Twitter’s term for the users that a given user follows) who also participated in the protest. Users whose follower or friend lists were set to private were excluded from analysis.

The third stage of data collection and construction was to develop a user profile for each user. In addition to the protest tweet data, I also collected each user’s past 200 tweets. From this data, I extracted the number of tweets that the user had posted that were related to the Black Lives Matter movement and the number of times each user was mentioned by another user in the network. Users whose tweets were protected or who had fewer than five total tweets were excluded from analysis.
Using each user’s Twitter history and information available from Twitter’s API, I was able to construct a profile of each user. The profile included:

1) **Number of followers**: the total number of followers each user had

2) **Number of BLM tweets**: the number of tweets that the user posted that were related to the Black Lives Matter movement in their last 200 tweets

3) **Number of Mentions**: the number of times that a user was mentioned by other users in the original protest network

4) **Activation Time**: the time of the user’s first tweet relating to the protest movement, measured in seconds from the time of the first tweet of the entire protest movement.

5) **Threshold**: the number of accounts that the user follows that had already joined the movement at the time of the user’s activation

6) **Eigenvector centrality**: a measure of network centrality that approximates a node’s influence in a graph based on its structural position in the network

Finally, in order to analyze the network dynamics of each protest, I also built an “influence” graph. I did this by adding an edge from node A to node B if 1) A follows B and 2) B has an activation time before A. The reasoning behind this construction supposes that A observes that B has joined the protest and is then influenced by B’s activation to join the protest. Thus, an edge from A to B can be coded as “node A is influenced by node B.” A visual representation of this model is presented in figure 1. The in-degree of each user corresponds to the number of other users that he or she directly influences to join the protest. The eigenvector centrality of each node on this influence graph is a measure of how much each user indirectly influences others to join the protest. Consider a case in which node A only influences a single node, B, to join the movement, but node B goes on to influence fifty other nodes to join. Node A
only directly influenced a single node to join the movement, but indirectly influenced fifty nodes to join the movement through his or her influence on node B. On the influence network created by the process described, this would be captured by node A having a low in-degree (in fact, an in-degree of 1), but a high eigenvector centrality score. This influence graph allowed me to measure two dependent variables:

1) **Number of followers directly influenced**: the number of a user’s followers that join the protest movement after the user joins

2) **Indirect influence score**: the eigenvector centrality of the user on the influence network described above, a proxy measure of indirect influence

*Figure 1*

Network visualizations and analysis were done with Gephi, R, and Matlab. Coding for the collection of the Twitter data was done in Python using the Twython package.

**Results and Discussion**

The results and discussion sections are split into two parts. The first part is a summary multivariate analysis of the four protest movements. The second part is an in-depth case study of the Andrew Loku protest.

**Summary analysis**
Basic network analysis was run for each protest event. Summary data of the protest events can be found in table 1. The protest events had an average number of participants of 604, with a standard deviation of 160.

*Table 1: Descriptive network statistics of each of the Black Lives Matters Twitter protests*

<table>
<thead>
<tr>
<th>Protest</th>
<th>Number of Tweets</th>
<th>Number of Participants</th>
<th>Avg. Clustering Coefficient</th>
<th>Avg. Degree</th>
<th>Avg. Path Length</th>
<th>Network Diameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andrew Loku</td>
<td>953</td>
<td>561</td>
<td>.35</td>
<td>19.78</td>
<td>3.15</td>
<td>9</td>
</tr>
<tr>
<td>Greg Gunn</td>
<td>779</td>
<td>428</td>
<td>.29</td>
<td>11.85</td>
<td>4.4</td>
<td>11</td>
</tr>
<tr>
<td>Akiel Denkins</td>
<td>1498</td>
<td>814</td>
<td>.185</td>
<td>9.8</td>
<td>5.12</td>
<td>14</td>
</tr>
<tr>
<td>Che Taylor</td>
<td>1086</td>
<td>614</td>
<td>.16</td>
<td>6.28</td>
<td>4.37</td>
<td>13</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>1079</strong></td>
<td><strong>604</strong></td>
<td><strong>0.25</strong></td>
<td><strong>11.93</strong></td>
<td><strong>4.26</strong></td>
<td><strong>12</strong></td>
</tr>
<tr>
<td><strong>Standard Deviation</strong></td>
<td><strong>306</strong></td>
<td><strong>160</strong></td>
<td><strong>.09</strong></td>
<td><strong>5.72</strong></td>
<td><strong>.82</strong></td>
<td><strong>2.2</strong></td>
</tr>
</tbody>
</table>

Multiple linear regression was used to test a model for predicting the number of followers directly influenced by each user from their number of mentions ($X_1$), number of BLM tweets ($X_2$), number of followers ($X_3$) and eigenvector centrality ($X_4$):

\[
\text{Number of followers directly influenced} = B_0 + B_1X_1 + B_2X_2 + B_3X_3 + B_4X_4 + \text{error}
\]

Table 2 reports the descriptive statistics of the variables in the analysis. Table 3 reports the results of the multiple linear regression model. The first column reports the unstandardized coefficient. The second column reports the standard error. The third column reports the standardized coefficient.

*Table 2: Descriptive statistics and descriptions of the variables used in the multivariate analysis*

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Description</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of mentions</td>
<td>The number of times a user was mentioned by other users in the network, including historical tweets</td>
<td>8.21</td>
<td>55.03</td>
</tr>
<tr>
<td>Number of BLM Tweets</td>
<td>The number of tweets related to the black lives matters movement the user has tweeted, including</td>
<td>12.40</td>
<td>20.47</td>
</tr>
</tbody>
</table>
historical tweets

<table>
<thead>
<tr>
<th>Number of followers</th>
<th>The total number of followers a user has</th>
<th>2211.03</th>
<th>13737.04</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eigenvector Centrality**</td>
<td>The eigenvector centrality of the user on the full network</td>
<td>.056</td>
<td>.11</td>
</tr>
</tbody>
</table>

**Dependent Variable**

| Number of followers directly influenced | The number of a user’s followers that join the protest movement after the user joins; a measure of direct influence | 5.39 | 20.31 |
| Indirect influence score | The eigenvector centrality of the user on the influence network described above; a measure of indirect influence | .017 | .064 |

**Only used as an independent variable for predicting number of followers directly influenced.**

Table 3: Summary of Multiple Regression Analysis for the Number of Followers Directly Influenced

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>SE</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-.931</td>
<td>.362</td>
<td></td>
</tr>
<tr>
<td>Number of mentions</td>
<td>.237</td>
<td>.007</td>
<td>.633**</td>
</tr>
<tr>
<td>Number of BLM Tweets</td>
<td>.033</td>
<td>.015</td>
<td>.033*</td>
</tr>
<tr>
<td>Number of Followers</td>
<td>-7.71E-5</td>
<td>.000</td>
<td>-.052**</td>
</tr>
<tr>
<td>Eigenvector Centrality</td>
<td>73.260</td>
<td>3.534</td>
<td>.402**</td>
</tr>
</tbody>
</table>

\( Adjusted R^2 = .80 \)
\( F(4, 1288) = 746.52 \)

* \( p < .05 \)
** \( p < .001 \).

Each of the predictor variables had significant effects. According to the results, higher mention count, eigenvector centrality, and Black Lives Matters tweet count each predicted higher numbers of followers influenced. Conversely, higher total number of followers predicted lower numbers of followers influenced. The model accounted for 80.0% of the variance in the number of followers influenced, \( F(4, 1288) = 746.52, p < .001, R^2 = .800. \)

A multiple linear regression model was also used to predict the indirect influence score of each user from their number of mentions \( (X_1) \), number of BLM tweets \( (X_2) \), and number of followers \( (X_3) \):

\[
\text{Indirect influence score} = B_0 + B_1X_1 + B_2X_2 + B_3X_3 + \text{error}
\]
Table 4 reports the results of the multiple linear regression model. Eigenvector centrality on the full network was not included as a predicting factor due to its similarity to the dependent variable.

Table 4: Summary of Multiple Regression Analysis for the Indirect Influence Score

<table>
<thead>
<tr>
<th>Variable</th>
<th>$B$</th>
<th>$SE$</th>
<th>$\beta$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>.006</td>
<td>.002</td>
<td>.724**</td>
<td>.724**</td>
</tr>
<tr>
<td>Number of mentions</td>
<td>.006</td>
<td>.000</td>
<td>.724**</td>
<td>.724**</td>
</tr>
<tr>
<td>Number of BLM Tweets</td>
<td>.001</td>
<td>.000</td>
<td>.119**</td>
<td>.119**</td>
</tr>
<tr>
<td>Number of Followers</td>
<td>0.000</td>
<td>.000</td>
<td>.038</td>
<td>.038</td>
</tr>
</tbody>
</table>

$Adjusted R^2 = .605$

$F(4, 1288) = 746.52$

*p < .05

**p < .001

Each of the predictor variables had a significant ($p < .05$) zero-order correlation with the number of followers recruited. However, only number of mentions and number of Black Lives Matters tweets were significant predictors of the indirect influence score. Number of followers was not a significant predictor. The model accounted for 60.5% of the variance in the indirect influence score, $F(4, 1288) = 284.4, p < .001, R^2 = .605$.

Summary Discussion

My analysis of protest data related to the Black Lives Matters movement revealed some expected as well as some surprising results. First, I found that the number of times a user is mentioned in the protest network is the strongest predictor of their direct and indirect influence in the network. Results show that a user with 10 more mentions will directly influence 2.37 more people to join the protest and have an indirect influence score that is .6 units higher. Users who
engage the attention of their followers were significantly more influential in recruiting other 
users to the protest movement, according to the model described in this paper.

Additionally, I found that users with higher eigenvector centrality scores significantly 
predicted higher levels of direct influence. The eigenvector centrality was almost as strong of a 
predictor of direct influence as the number of mentions; eigenvector centrality had a standardized 
coefficient of .402 while number of mentions had a standardized coefficient of .633. According 
to the results, a user with an eigenvector centrality of .1 units higher will directly influence 7.3 
users to join the protest. This finding suggests that users with a more central network position 
recruited more of their followers to join the protest movement than those on the outskirts of the 
network.

I also found the number of Black Lives Matters tweets a user posted predicted 
significantly higher levels of direct and indirect influence in the protest network, although this 
factor was not as strong of a predictor as the number of mentions and the eigenvector centrality. 
A user that has 30 more tweets about the Black Lives Matters movement will directly influence 1 
more user to join the protest. Still, this finding suggests that prior involvement in the Black 
Lives Matters movement on Twitter predicts higher levels of influence in a new Black Lives 
Matters protest.

The results, however, did not support the hypothesis that the more followers a user has, 
the more users he or she will directly or indirectly influence to join a protest movement. My 
analysis initially showed the number of followers a user has to be weakly positively correlated 
with the number of users they directly and indirectly influenced. However, when I performed a 
multivariable regression with other controlling factors, I found that higher number of followers 
actually predicted lower numbers of users influenced, the exact opposite of the effect we
expected. While the effect is significant, it is very small; an increase in 15000 followers will directly influence 1.05 fewer users to join the protest. Thus, while users with more followers tend to influence more people to join the movement, their influence is not because they have more followers. Rather, having more followers is significantly correlated with being in a more central network position and attracting higher numbers of mentions, both factors that significantly lead to higher levels of influence in the network. The number of followers was only serving as a proxy for these true predictors of influence.

It is possible that I found this effect because users with the largest numbers of followers tend to be celebrities, brand accounts, or news accounts whose followers do not necessarily have a personal connection with the user. In these cases, the follower relationship might constitute a weak tie. Research has shown that while weak ties are crucial in processes of contagion including the spread of disease, adoption of new technologies, and the diffusion of job information, they are less relevant for social contagion of higher risk behavior such as protest diffusion (Centola & Macy, 2007). In these cases, strong ties are more successful for spreading social behavior.

The importance of strong ties might shed light on the strong relationship between the number of times a user was mentioned by the protest network and how influential the user was in the network. On Twitter, the number of times a user engages another user through mentions could be an indication of the strength of the tie between users. Thus, users who attract a lot of mentions have stronger ties with their followers, which might explain their higher levels of influence in the network.

Prior studies have already shown found that while users with many followers tend to be more successful at triggering information cascades, they are generally unreliable predictors
(Bakshy et al., 2011). Simply having many followers does not mean that one is particularly successful at engaging or influencing those followers. My study went a step further and showed that number of followers is an inverse predictor of ability to spread activity related to protest movements. An explanation for this difference might be that the previous study mentioned did not look at the dissemination of protest related tweets in particular. Users with many followers might be more relatively influential when spreading information about less controversial subjects. However, when it comes to protest recruitment particularly, stronger ties matter more.

**Case Study Analysis of the Andrew Loku Protest**

Tweets related to the Andrew Loku protest were collected for the week of March 18th, 2016 to March 24th, 2016. Most activity was in response to the lack of charges brought against the Toronto police officer who shot an unarmed black man, Andrew Loku. In total, 953 tweets were collected from 561 different users. Figure 2 shows a network visualization of the graph, with the nodes colored by modularity class and sized by eigenvector centrality. The visualization shows that there are two separate clusters roughly corresponding to the Toronto BLM network (green) and the national BLM network (orange). Figure 3 shows the number of users who joined the protest as a function of time for the week of March 18th to March 24th, where time is given in seconds since the start of the movement. Figure 4 shows the clustering coefficient of the network, a measure of the degree to which the nodes of a network cluster together, a function of time. The figure shows that the clustering coefficient of the network doubles around timestep 180,000, with the network growing more clustered over time. Figure 5 shows the threshold to participate of each user, where participation is defined as tweeting using one of the hashtags related to the protest movement and threshold is defined as the number of users who the user follows who had joined the protest at the time of the user chose to participate. The average
threshold for all users was 5.39 ($N=561$, $SD=5.61$). Finally, a table of the users with the highest levels of eigenvector centrality, between-ness centrality, direct influence, and indirect influence can be found in Tables 5-8, respectively.

*Figure 2: Network visualization of the Andrew Loku Twitter protest. Nodes are sized by eigenvector centrality and colored by modularity class.*
Figure 3: The number of users who joined the Andrew Loku protest as a function of time. Time is given in seconds since the start of the protest.

![# Nodes Time Series](image)

Figure 4: The clustering coefficient of the Andrew Loku protest as a function of time. Time is given in seconds since the start of the protest.

![Clustering Coefficient Time Series](image)
Figure 5: Threshold curve for the #justice4andrewloku protest. \( P(K) \) is the fraction of users who joined the protest after their \( K^{th} \) friend joined the protest.

Table 5: Descriptions of the users in the Andrew Loku protest with the highest eigenvector centrality

<table>
<thead>
<tr>
<th>Rank</th>
<th>Username</th>
<th>Eigenvector Centrality, Full</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DesmondCole</td>
<td>1.0</td>
<td>Personal account, Toronto journalist</td>
</tr>
<tr>
<td>2</td>
<td>BLM_TO</td>
<td>.981</td>
<td>Official account of BLM Toronto</td>
</tr>
<tr>
<td>3</td>
<td>Blklivesmatter</td>
<td>.565</td>
<td>Official account of BLM movement</td>
</tr>
<tr>
<td>4</td>
<td>judyrebick</td>
<td>.538</td>
<td>Personal account, Toronto journalist</td>
</tr>
<tr>
<td>5</td>
<td>metromorning</td>
<td>.533</td>
<td>Media account, Toronto radio station</td>
</tr>
</tbody>
</table>

Table 6: Descriptions of the users in the Andrew Loku protest with the highest betweenness centrality

<table>
<thead>
<tr>
<th>Rank</th>
<th>Username</th>
<th>Betweenness Centrality</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DesmondCole</td>
<td>.212</td>
<td>Personal account, Toronto journalist</td>
</tr>
<tr>
<td>2</td>
<td>BLM_TO</td>
<td>.122</td>
<td>Official account of BLM Toronto</td>
</tr>
<tr>
<td>3</td>
<td>AnthonyNMorgan</td>
<td>.052</td>
<td>Personal account, Toronto Activist</td>
</tr>
<tr>
<td>4</td>
<td>OccupyToronto</td>
<td>.041</td>
<td>Official account of the Toronto Occupy movement</td>
</tr>
<tr>
<td>5</td>
<td>IdleNoMore4</td>
<td>.034</td>
<td>Media account, Toronto radio station</td>
</tr>
</tbody>
</table>
Table 7: Descriptions of the users in the Andrew Loku protest who influenced the most number of followers to join the protest

<table>
<thead>
<tr>
<th>Rank</th>
<th>Username</th>
<th>Number of followers influenced</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>BLM_TO</td>
<td>340</td>
<td>Official account of BLM Toronto</td>
</tr>
<tr>
<td>2</td>
<td>DesmondCole</td>
<td>228</td>
<td>Personal Account, Toronto Journalist</td>
</tr>
<tr>
<td>3</td>
<td>Blklivesmatter</td>
<td>140</td>
<td>Official account of the BLM movement</td>
</tr>
<tr>
<td>4</td>
<td>prisonculture</td>
<td>81</td>
<td>Personal account, unofficial activist</td>
</tr>
<tr>
<td>5</td>
<td>judyrebick</td>
<td>73</td>
<td>Personal account, Toronto journalist</td>
</tr>
</tbody>
</table>

Table 8: Descriptions of the users in the Andrew Loku protest with the highest indirect influence scores

<table>
<thead>
<tr>
<th>Rank</th>
<th>Username</th>
<th>Indirect Influence Scores</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>BLM_TO</td>
<td>1</td>
<td>Official account of BLM Toronto</td>
</tr>
<tr>
<td>2</td>
<td>DesmondCole</td>
<td>0.455054181</td>
<td>Personal account, Toronto journalist</td>
</tr>
<tr>
<td>3</td>
<td>AnthonyNMorgan</td>
<td>0.450530486</td>
<td>Personal account, unofficial activist</td>
</tr>
<tr>
<td>4</td>
<td>Blklivesmatter</td>
<td>0.419559445</td>
<td>Official account of the BLM movement</td>
</tr>
<tr>
<td>5</td>
<td>prisonculture</td>
<td>0.369077696</td>
<td>Personal account, unofficial activist</td>
</tr>
</tbody>
</table>

Case Study Discussion

A case study analysis of the Andrew Loku provides a more holistic picture of this particular Black Lives Matters protest and the way in which it unfolded on Twitter. Figure 3 shows the number of users who had tweeted about the movement from March 18th, 2016 to March 24th, 2016. A regular user (whose name is omitted for privacy reasons) sent the first tweet of the movement on March 18th, but the movement appears to jump in participation later on March 18th after @BLM_TO first tweeted. Periods of steady recruitment are punctuated by periods in which large numbers of users join within a short window of time. These spikes in recruitment seem to follow the activation of influential users, although a quantitative analysis was not performed.

Additionally, I identified a list of the users who were the most influential according to the measures of direct and indirect influence constructed in this paper. I also identified a list of nodes
that were high in between-ness centrality and eigenvector centrality. The centrality measures of the network show that largely journalists, media outlets, and official accounts of the Black Lives Matters movement were most central in the movement. However, the influence measures constructed in this paper also show that activists without any “official” role in the BLM movement were influential in the network, in addition to journalists and official accounts of the Black Lives Matters Movement. The importance of official accounts and journalists show the ways in which the Black Lives Matters Movement has begun to organize formally and gain the attention of mainstream media as it has matured, while the continued importance of unofficial activists demonstrates the way in which it is still a grassroots movement.

That being said, @BLM_TO and @BlkLivesMatters were particularly important in driving protest activity. @BLM_TO is the official Twitter account of the Black Lives Matters movement in Toronto, where Andrew Loku was killed, and @BlkLivesMatters is the official Twitter account of the international Black Lives Matters movement. Analysis of the clustering coefficient of the graph over time seems to shed more light on how these official accounts drive movement participation. After @BLM_TO tweets on March 18th the clustering coefficient of the graph remains fairly constant until March 21st and then doubles soon after the official @BlkLivesMatters Twitter account first tweets about the movement. Coloring the nodes in figure 2 by modularity class shows that the network is grouped into two clusters with those two accounts at the center of each. The green cluster seems to correspond roughly to the Toronto participants while the orange cluster seems to correspond to the national Black Lives Matters community. While obviously a limited sample, the existence of these separate communities sheds light on the organization of this movement. In this case at least, there seems to be a community of activists in Toronto who are driving protest activity there. This local clustering of
Twitter users led by the official Black Lives Matter Twitter account reflects the chapter-like structure of the organization and demonstrates a higher level of organization than one might expect from a purely grassroots campaign.

Finally, the threshold curve of the users sheds light on the way in which users are recruited to the movement. Interestingly, a fifth of users had a threshold of one, meaning only one of the people they followed had tweeted about the protest at the time of their first tweet. The user threshold curve is more skewed towards lower values than in other literature on protest recruitment, even other protest recruitment on Twitter (Gónzalez-Bailon et al, 2011; Romero, Meeder, & Kleinberg et al). In general, the threshold for participants in this protest seems lower than expected based on the literature. There are several possible explanations for this effect. First, it is possible that users with higher thresholds were simply not activated, either because the window was too short or because not enough users joined to activate higher threshold users. Second, it is likely that the lower costs to participation of joining a protest on Twitter rather than an in person protests lowers the thresholds of participation Twitter users. Finally, it could be the case that the users in the network are already familiar with the Black Lives Matters movement and are more willing to join protests related to the cause than other more unfamiliar causes. However, without being able to compare this data to other protests or topic diffusions on Twitter, one cannot draw definitive conclusions. Methods to further examine the possible reasons for this effect are elaborated upon in the “Directions for Further Research” section.

Conclusion

In summary, our analysis of Black Lives Matters Twitter protests shed light on which users are most influential in the BLM movement and how these protests unfolded on Twitter. We
also identified ways in which this protest both conformed to and defied expectations drawn from previous literature on protest recruitment.

In particular, we found that users who engaged their followers more, had an advantageous network position, and were more active in the Black Lives Matters community on Twitter were the most influential at recruiting other participants to protest activity. Most often, these actors were journalists and official Black Lives Matters Twitter accounts, although some were activists not officially affiliated with the BLM movement. While many of these influential users had large numbers of followers, merely having large numbers of followers absent these other factors was not advantageous in recruiting other users.

This analysis shed light on the way that a more mature protest movement functions. Like previous Twitter movements, users who engage their followers more are more influential at recruiting participants. However, unlike other Twitter protests that were studied at their inception, media and official movement accounts played a large role in driving protest activity and recruiting participants, reflecting the top-down organization of the movement in its later stage.

Despite the wealth of these findings, there is room for further analysis of protest recruitment and information dissemination on Twitter. Some of these avenues, as well as limitations of this current study, are explored in the next section.

**Limitations and Directions for Further Research**

One drawback of this study is that most data collected related to the protests was for a very short time window (approx. one week per protest) due to Twitter’s rate limits. A more comprehensive analysis of this movement would look at the evolution of protest dynamics and recruitment over a longer period of time. One area of interest would be to compare Black Lives
Matters Twitter activity at its inception to activity now to examine how the movement evolved over time. It would be interesting to study who the key players were at the beginning of the movement and to look at whether and how that has changed. Such analysis would likely be possible with the cooperation of Twitter or through the purchase of tweets through a third party company.

Additionally, a further area for research would be to compare how protest activity on Twitter differs from other activity on Twitter. It would be interesting to compare protest diffusion on Twitter to the diffusion of other trending topics. My analysis of the thresholds of protesters suggests that there might be an active Black Lives Matters community on Twitter that allows for swift recruitment to new protests, but it also might be the case that other less controversial or less political memes might disseminate more quickly. A comparative analysis would shed light on the subject.

There are many other directions for further research beyond the ones laid out in this paper. Twitter and other social media platforms provide easy access to a wealth of data related to protest recruitment and information dissemination. The world has shifted the way it shares information and organizes collective action, and research should evolve to reflect that as well.
References


