Is Slacktivism Underrated? Measuring the Value of Slacktivists for Online Social Movements

Lia Bozarth and Ceren Budak
University of Michigan, School of Information
105 S State St, Ann Arbor, MI 48109
{lbozarth, cbudak}@umich.com

Abstract

We measure the value of "slacktivists" involved in more than 80 protest movement hashtags on Twitter, focusing on their value for content generation, communication, encouragement, and content diversity. Further, we move beyond the existing literature which primarily focuses on direct content contributions by investigating whether slacktivists induce further contributions from other activists by providing them simple cues of encouragement that enhance non-slacktivists' commitment to a movement. Our findings suggest that slacktivist actions, unlike those by dedicated activists, do not encourage others, and if anything, discourage them.

Introduction

In the era of Web 2.0, online activism often serves as the catalyst for revolutionary protests and demonstrations (Radsch 2011; Newsom and Lengel 2012). Yet, skeptics refer to it as *slacktivism* (Morozov 2009), arguing that these simple online actions neither persuade nor promote real change (Lewis, Gray, and Meierhenrich 2014). Proponents, however, contend that online activism is valuable in its ability to raise and force issue awareness for a broad populace (González-Bailón, Borge-Holthoefer, and Moreno 2013).

Most studies treat all online activists interchangeably; nevertheless, their individual effort can vary widely. We make this distinction in examining the lowest effort among low-effort individuals: those whose whole contribution to an online movement consists of a single tweet. We evaluate these slacktivists based on more than 80 race or gender-related hashtags with respect to content, conversation, and encouragement volumes, and content diversity.

Coleman (1988) posited that social incentive provided by free-riders may encourage zealots to sustain their participation. Inspired by this theoretical study, we estimate indirect contributions by slacktivists in the form of social incentive to others. Our analyses surprisingly show that slacktivists are associated with a *negative* effect. Individuals encouraged by slacktivists are not only less likely to continue their participation compared to those who are encouraged by non-slacktivists, they are also less likely to participate further than those who were not encouraged by anyone at all.

Related Work

Is online activism a legitimate form of activism with real impact? Many researchers say yes. Starbird et al. (2012), focusing on the 2011 Egyptian revolution, argued that online activism provides value by expressing solidarity and processing information. Tufekci et al. (2012) concluded that online activism was crucial in influencing individual decisions about protest participation. Gleason (2013) made a similar argument in studying the Occupy Wall Street protests. Gonzalez-Bailon et al. (2013), undertaking study of the indignados movement, emphasized online activism's importance for fast diffusion of ideas and organizational details. And Barbera et al. (2015) asserted that while individual slacktivists contribute little, their aggregated contribution rivals that of the core zealots.

Alternatively, Lewis et al. (2014), studying the "Save Darfur" campaign on Facebook, argued that online activism generates surprisingly little "real" activism as measured by donations or recruitment of new members. Morozov (2009) went so far as to argue that slacktivism hurts traditional activism since ordinary people would prefer the lazier solution.

Motivated to provide further insights, we focus on a particularly low-cost form of online activism: tweeting a single tweet, and measure its value for online protest movements.

This paper is also related to a number of recent papers on social movements examined through the lens of social media (González-Bailón et al. 2011; Earl and Kimport 2011; Starbird and Palen 2012; González-Bailón, Borge-Holthoefer, and Moreno 2013; Hanna 2013; Weber, Garimella, and Batayneh 2013; Budak and Watts 2015; Barberá et al. 2015; De Choudhury et al. 2016; Freelon, McIlwain, and Clark 2016). Unlike such studies that focus on a single or few movements, here we study more than 80 protest hashtags varying in topic (race or gender) and size.

Data

Our research dataset consists of 49.5 million tweets from 7.3 million distinct Twitter users from Feb. 1, 2014 to May 10, 2015, spanning 49 race-related and 36 gender-related hashtags. Among these tweets, 16.3 million are replies that involve 11.6 million unique pairs of Twitter users. Similarly, the set includes 36.1 million retweets that involve 27 million unique pairs of Twitter users. The list of the protest hashtags and some high level statistics about each protest hashtag are provided in *http://cbudak.com/projects/slacktivists.html*.

The race-related hashtags cover the stream of demonstrations and riots in Baltimore, Ferguson, and New York City, among others, against the systemic racism and violence toward African Americans within the criminal justice system and the police killing of Eric Garner, Freddie Gray, Michael Brown, Tamir Rice, et cetera. The gender-related hashtags involve separate but related issues such as sexism in the news and entertainment media, women's health care, women's career and leadership, men's role in feminism, and rape culture.

Evaluating Slacktivism

In this section, we describe how we measure the contributions of slacktivists in online activism. First, we list important definitions the rest of the manuscript relies on.

Social Movement: A movement, m_i, is defined as content that includes a hashtag, $\#m_i$. Although certain hashtags are extensions belonging to the same umbrella protest, treating them as separate movements allows for a more precise analysis of their similarities and differences.

Slacktivist: Slacktivists for a movement m_i are defined as S_i, the set of users who tweeted at most one tweet with a hashtag, $\#m_i$. Note that slacktivists are defined at a per-hashtag level; that is, an individual user who posted a single tweet using a hashtag m_i is defined as a slacktivist for m_i, irrespective of how much she contributes to other movements. Furthermore, a tweet containing multiple hashtags is counted once for each movement it belongs to.

Direct Contributions: We address the following direct contributions in this paper.

1. Tweet volume: We measure the fraction of content in m_i that is generated by slacktivists, or S_i.

2. Endorsement volume: We measure a) the fraction of endorsements (retweets) generated by S_i, and b) the fraction of endorsements presented to S_i (retweets of slacktivists' tweets by others).

3. Conversational volume: We measure a) the fraction of conversations (mentions or @messages) generated by S_i, and the fraction of conversations presented to S_i (mentions of slacktivists by others).

We selected endorsement and conversational measures based on previous research by Chen (2011), who suggested that these elements gratify users' need for connection.

4. Content diversity: Kullback-Leibler (KL) divergence is commonly used to determine dissimilarity between text documents (Huang 2008). Here, we use it to estimate information loss if the corpus of the content contributed by slacktivists is used to estimate that by the non-slacktivists, and vice versa. We do this is to examine whether content from the two groups is interchangeable.

Induced Contributions: Coleman (1988) theorized that positive external intervening activities from an individual's social network are capable of inducing zealotry in that individual. Following this logic, we examine whether being endorsed by or being engaged in a conversation with a slacktivist increases the likelihood of non-slacktivists to continually participate in online protests.

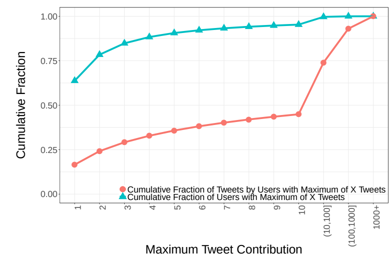


Figure 1: Cumulative Contribution Overview.

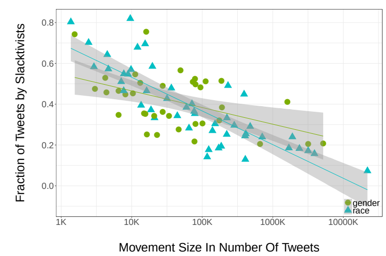


Figure 2: Slacktivist Contribution and Movement Size.

Results

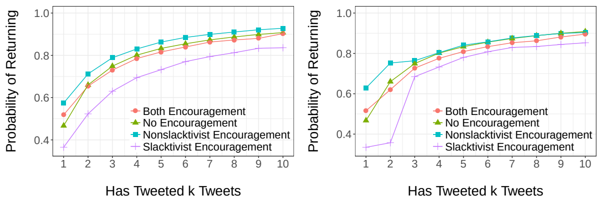
Direct Contributions We review tweet, endorsement and conversation volumes, and content diversity in detail.

Tweet volume: As shown in Figure 1, there is a considerable variance in the number of tweets a single Twitter user contributes–a large number of people contribute very little while a smaller fraction are highly dedicated. We observe that on an aggregated level, S accounts for 23.3% of all the tweets and 63.7% of all the users.

In Figure 2 we demonstrate the variance across different movements for S_i. Here, the x-axis denotes the size of the movement (total number of tweets in log scale), and the y-axis denotes the fraction of tweets in a movement generated by the slacktivists. The shape of each data point indicates whether it is a race or gender related movement. Unsurprisingly, we observe a clear trend that smaller movements have a higher fraction of tweets generated by slacktivists. Globally, the slope for gender-related hashtags is flatter than the slope for race-related hashtags–size matters less in defining the content contribution value of slacktivists for gender related hashtag movements.

Endorsement and conversation volumes: Slacktivists account for 29.6% of all retweets and 13.7% of all mentions. Furthermore, retweets of slacktivists' tweets account for 13.5% of all retweets, and mentions of slacktivists account for 6.5% of all mentions. We also examine how these four dimensions interact with protest size. The results are similar to those seen in Figure 2; therefore we omitted them here due to space limitations. Overall, the findings show that the slacktivists who individually contribute only a single tweet collectively make up a large fraction of movements.

Content diversity: For each hashtag, we group tweets by slacktivists into a single document and calculate its word distribution as W_{sl}; we employ the same procedure to get



(a) Retweets (b) Mentions

Figure 3: Inductive Effect of Encouragement on the Probability of Participants Coming Back for More Activity. Encouragement is defined in terms of retweets in (a), and in terms of mentions in (b). Ninety-five percent confidence intervals are also plotted but are not visible given the small interval.

non-slacktivist word distribution W_{ns}. We then apply KL divergence to show the amount of information loss if content from slacktivists is used to approximate content from non-slacktivists using the formula:

$$D_{KL}(W_{ns}||W_{sl}) = \sum_i W_{ns}(i) log \frac{W_{ns}(i)}{W_{sl}(i)} \tag{1}$$

We re-apply the equation with the positions of W_{ns} and W_{sl} switched to calculate information loss in reverse (using non-slacktivist content to approximate slacktivist content). We observe that information loss is roughly two times greater when content from slacktivists is used to approximate content from non-slacktivists for 95.2% of the hashtags. This suggests that slacktivist tweets have relatively lower information entropy, or are less interesting. We also observe a marked difference in information loss for different groups of protests; results are summarized in Table 1. We see, for instance, that the word distribution of slacktivist and non-slacktivist content differs more for gender-related movements as opposed to race-related movements. We can also see that average information loss is highest for the smallest hashtags and decrease as movement size increases. However, it's noteworthy that not all hashtags adhere to the same pattern. Take #iftheygunnedmedown for example, information loss from using non-slacktivists' content to approximate slacktivists' content is actually greater than the other way around despite the hashtag's considerable size.

Table 1: KL Divergence Summary. Note that $\mu_{W_{ns}||W_{sl}}$ is the mean KL score when content from slacktivists is used to approximate content from non-slacktivists; $\sigma_{W_{ns}||W_{sl}}$ is the standard deviation. Higher μ indicates more information loss.

Group	$\mu_{w_{ns}\|\|w_{sl}}$	$\sigma_{w_{ns}\|\|w_{sl}}$	$\mu_{w_{sl}\|\|w_{ns}}$	$\sigma_{w_{sl}\|\|w_{ns}}$
By hashtag type				
Race	1.25	0.01	0.68	0.01
Gender	1.79	0.04	0.83	0.02
By movement size				
$[10^3, 10^4)$	2.20	0.08	1.19	0.04
$[10^4, 10^5)$	1.63	0.02	0.80	0.02
$[10^5, 10^6)$	0.89	0.01	0.43	0.01
$[10^6, \infty)$	0.64	0.02	0.26	0.01

Induced Contributions Next, we investigate whether slacktivists further contribute to these movements through induced contributions. We start by applying a simple regression test to determine whether individuals who are retweeted or mentioned (irrespective of slacktivist or non-slacktivist standing) are significantly more likely to come back and make further contributions to a movement. The results show they are, true across all movements (omitted due to space limitations). Yet, the following question remains: Is there a difference in the effect of an endorsement between a slacktivist and non-slacktivist?

To answer the question, we perform the following analysis: Let $U^{retweet}_{sl,k}$, $U^{retweet}_{ns,k}$, $U^{retweet}_{b,k}$, $U^{retweet}_{n,k}$ be the set of users who tweeted k tweets across all movements and were retweeted by 1) at least one slacktivist but no non-slacktivist, 2) at least one non-slacktivist but no slacktivist, 3) at least one slacktivist and one non-slacktivist, and 4) no one after their k^{th} tweet. We denote the probabilities that a user in these groups would came back to tweet at least once more after their k^{th} tweet as $p^{retweet}_{sl,k}$, $p^{retweet}_{ns,k}$, $p^{retweet}_{b,k}$, $p^{retweet}_{n,k}$. Comparing these four probabilities can help us understand whether retweets from slacktivists and non-slacktivists are associated with higher rates of retention and if so, determine which one has a higher effect.

The results are in Figure 3(a) and show a surprising pattern. As expected, we see that the non-slacktivists are associated with a positive effect size–those retweeted by a non-slacktivist are significantly more likely to come back than those who are not retweeted by anyone. The effect size is notable. For instance, for users who have only one tweet in the movements (x=1), the likelihood to come back is 46.7% for those with no retweets and 56.1% for those who are retweeted by at least one non-slacktivist—a 20.1% improvement. The effect size diminishes for larger k. For individuals with an already high investment in a movement (e.g., already having 10 tweets), the difference between being retweeted by a non-slacktivist and by no one is much smaller. The pattern for those who are retweeted by a slacktivist is surprisingly different. Unlike what would be predicted through Coleman's theory, we observe that the encouragements provided by slacktivists are associated with a negative effect. Those who are retweeted by a slacktivist are not only less likely to come back compared to those retweeted by a non-slacktivist–but they are less likely to come back even when compared to those who have not been retweeted by anyone at all. In addition, unlike the case of non-slacktivists, the effect size is rather consistent for the different values of k, i.e. for individuals with different levels of demonstrated dedication to the movement. The pattern observed for those who are retweeted by both a slacktivist and a non-slacktivist is more positive than the slacktivist case and more negative than the non-slacktivist case.

Next, we study the effect of *communication* as an incentive to sustain movement participant retention. Similar to the case of retweets, we define $U^{mention}_{sl,k}$, $U^{mention}_{ns,k}$, $U^{mention}_{b,k}$, $U^{mention}_{n,k}$ as the set of users who tweeted k tweets in all movements and were *mentioned* by 1) at least one slacktivist but no non-slacktivist, 2) at least one non-slacktivist but no slacktivist, 3) at least one slacktivist and one non-slacktivist, and 4) no one after their k^{th} tweet. We estimate the likelihood of such individuals to come back to make further contributions to the movement. In Figure 3(b), we present the findings for $p_{sl,k}^{mention}$, $p_{ns,k}^{mention}$, $p_{b,k}^{mention}$, and $p_{n,k}^{mention}$. The results are largely consistent with the retweet findings. Those who are engaged in a conversation by at least one non-slacktivist are more likely to come back and make further contributions to a movement. This effect is significant and large for low values of k. Yet, unlike the case of retweets, the significance disappears beyond k=3. Again, surprisingly, those who are mentioned by a slacktivist are less likely to come back to make further contributions than their counterparts who have not been engaged by anyone in the movement in a conversation after their k^{th} tweet. This effect is particularly high for low k values but stays in significant levels even for larger k values.

Discussion

Our results suggest that the value of the low-effort slacktivists who protested by contributing only a single tweet is rather complex. On the plus side, these slacktivists make up a substantial fraction of the population and account for a substantial fraction of concrete text content irrespective of whether it is tweet, endorsement, or conversation. Conversely, the correlation between being engaged by slacktivists and a lower probability of additional participation suggests that social incentive provided by slacktivists fell short of sustaining and motivating the existing population engaged in online movements. One potential explanation is that an acknowledgment by a slacktivist might register more of a concern for privacy. It is also possible that slacktivist retweets might suggest to individuals that their actions have already accomplished an important task by bringing a new participant to the protest.

For future work, it's worth investigating whether the observed negative correlation is causal. Our initial efforts in building comprehensive models that account for confounds such as overall Twitter activity and popularity (omitted due to space limitations) indicate that the negative effect persists, yet there is still more to do. A potential in-the-wild experiment could create zealot and slacktivist bots and randomly select other participants to retweet. Because treatment would be fully at random, causality could be established. Observed differences in race and gender hashtags in this study also demonstrate a need for more detailed examination.

Last, note that our research is limited to Twitter which does not capture the full scope of online user participation, and that our data only pertain to the protest actions taken with a specific hashtag.

References

Barberá, P.; Wang, N.; Bonneau, R.; Jost, J. T.; Nagler, J.; Tucker, J.; and González-Bailón, S. 2015. The critical periphery in the growth of social protests. *PloS one* 10(11):e0143611.

Budak, C., and Watts, D. J. 2015. Dissecting the spirit of gezi: Influence vs. selection in the occupy gezi movement. *Sociological Science* 2:370–397.

Chen, G. M. 2011. Tweet this: A uses and gratifications perspective on how active twitter use gratifies a need to connect with others. *Computers in Human Behavior* 27(2):755–762.

Coleman, J. S. 1988. Free riders and zealots: The role of social networks. *Sociological Theory* 6(1):52–57.

De Choudhury, M.; Jhaver, S.; Sugar, B.; and Weber, I. 2016. Social media participation in an activist movement for racial equality. In *Tenth International AAAI Conference on Web and Social Media*.

Earl, J., and Kimport, K. 2011. *Digitally enabled social change: Activism in the Internet age*. Mit Press.

Freelon, D.; McIlwain, C. D.; and Clark, M. D. 2016. Beyond the hashtags:# ferguson,# blacklivesmatter, and the online struggle for offline justice.

Gleason, B. 2013. # occupy wall street: Exploring informal learning about a social movement on twitter. *American Behavioral Scientist* 0002764213479372.

González-Bailón, S.; Borge-Holthoefer, J.; Rivero, A.; and Moreno, Y. 2011. The dynamics of protest recruitment through an online network. *Scientific reports* 1.

González-Bailón, S.; Borge-Holthoefer, J.; and Moreno, Y. 2013. Broadcasters and hidden influentials in online protest diffusion. *American Behavioral Scientist* 0002764213479371.

Hanna, A. 2013. Computer-aided content analysis of digitally enabled movements. *Mobilization: An International Quarterly* 18(4):367–388.

Huang, A. 2008. Similarity measures for text document clustering. In *Proceedings of the sixth new zealand computer science research student conference (NZCSRSC2008), Christchurch, New Zealand*, 49–56.

Lewis, K.; Gray, K.; and Meierhenrich, J. 2014. The structure of online activism. *Sociological Science* 1:1–9.

Morozov, E. 2009. The brave new world of slacktivism. *Foreign Policy* 19(05).

Newsom, V. A., and Lengel, L. 2012. Arab women, social media, and the arab spring: Applying the framework of digital reflexivity to analyze gender and online activism. *Journal of International Women's Studies* 13(5):31.

Radsch, C. C. 2011. The revolutions will be hashtagged: Twitter turns 5 as the middle east demands democracy. *Huffington Post* 29.

Starbird, K., and Palen, L. 2012. (how) will the revolution be retweeted?: information diffusion and the 2011 egyptian uprising. In *Proceedings of the acm 2012 conference on computer supported cooperative work*, 7–16. ACM.

Tufekci, Z., and Wilson, C. 2012. Social media and the decision to participate in political protest: Observations from tahrir square. *Journal of Communication* 62(2):363–379.

Weber, I.; Garimella, V. R. K.; and Batayneh, A. 2013. Secular vs. islamist polarization in egypt on twitter. In *Advances in Social Networks Analysis and Mining 2013, ASONAM '13*, 290–297. ACM.