

On User Engagement across Social Media Campaigns to Curb Gender-based Violence

Prakruthi Karuna¹, Hemant Purohit¹, Bonnie Stabile², and Angela Hattery³

¹ Humanitarian & Social Informatics Lab, George Mason University, USA

² Schar School of Policy & Government, George Mason University, USA

³ Women & Gender Studies Program, George Mason University, USA
{pkaruna,hpurohit,bstabile,ahattery}@gmu.edu

Abstract. Gender-based violence (GBV) is a human-generated crisis, existing in various forms, including physical and sexual violence offline, and now online via harassment and trolling. To mitigate this crisis, various social media campaigns have emerged in recent times. While studying social media campaigns for different domains such as public health and natural crises has received significant attention in the literature, such studies for GBV are still in nascent form. We present the first preliminary study to examine engagement in such public anti-GBV campaigns using social media at the launch of one such campaign, with an eye to answering the following research questions: (1) Which types of users (e.g., organizations or individuals) engage in and across such campaigns? (2) Does campaign engagement vary by gender? This analysis provides a direction to further study dynamics of anti-GBV campaigns and interactions among participating individuals and organizations.

Keywords: Gender-based Violence, Humanitarian Informatics, Social Media, Campaigns, #ItsOnUs, #StateOfWomen, #HeForShe

Introduction and Related Work

Gender-based violence (GBV) encompasses ‘acts of violence ranging from online harassment to domestic assault and human trafficking’ [1], with women as primary targets. GBV “results in, or is likely to result in, physical, sexual or psychological harm or suffering to women” - United Nations (UN), 1995 [2]. According to the UN entity for gender equality and the empowerment of women - *UN Women* [1], globally 1 in 3 women experience physical or sexual violence. In the United States (US) alone, it is estimated that nearly - “1 in 5 women are raped,” and “1 in 4 women experience severe physical violence by an intimate partner” at some point during their lifetimes [5]. The American Association of Universities’ Campus Climate Survey reports a similar pattern of sexual assault - 23 percent - at institutions of higher education in the US [6]. Governmental and non-governmental (NGO) organizations from local to global levels have risen to combat the problem of GBV. To understand the institutional anti-GBV campaign’s efforts in the social media context, we investigate the aftermath of

the launch of an anti-GBV campaign *#StateOfWomen* to study whether such a public campaign brings awareness of other similar cause-driven campaigns in the social media discussion. *#StateOfWomen* is a national initiative, sponsored by the US White House, one facet of which focuses on violence against women in the US [4]. We observe the interaction of this campaign’s community (a group of social media users who use the campaign’s identity hashtag in the shared messages for discussion about this campaign) with two other closely-relevant and popular communities of anti-GBV campaigns – college-centric *#ItsOnUs* launched by the US White House[3] and global *#HeForShe* launched by the UN [1]. Our ultimate goal is to understand who engages in such campaigns, and how cause-driven campaigns play out in social media venues like Twitter. We seek to understand how such campaigns shapes people’s engagement towards the human-generated crisis of GBV, with a focus on understanding the reach and overlap of such anti-GBV campaigns with other relevant campaigns that share similar motives in social media.

Although there are several studies on social media for user engagement using hashtags in brand campaigns and activism [8, 7], there is a lack of literature on engagement in awareness campaigns for social issues in the GBV domain, and the recent few studies [10, 9] have only explored topical patterns for the overall GBV theme, while not focusing on online campaign dynamics of related causes for the use of nonprofits. Sociological Analysis of offline anti-GBV activism [11] showed that early anti-violence activists understood clearly the connections between GBV and other forms of injustice against women, including unequal pay and reproductive justice, and therefore, activism was driven by underlying related efforts of similar causes. Thus, more intensive domain-specific analysis of online interactions and collaboration among actively engaged users requires first a preliminary investigation of who participates and in which capacity across similar GBV campaigns. Given the adoption of Twitter as a major vehicle for awareness by NGOs and nonprofits in recent times [12], we study the following questions for the aforementioned three campaigns on Twitter:

- R1. Which types of users (e.g., organizations or individuals) engage in an anti-GBV campaign activism and is this activism observed across other similar campaigns?
- R2. Does anti-GBV campaign engagement vary by demographics such as gender?

Data Collection and Processing Method

We adopt a keyword-based crawling method for collecting Twitter messages (*tweets*) that is a common method for Twitter studies in the literature. Using Twitter Streaming API’s ‘filter/track’ method, which provides a stream of public tweets containing any of the provided seed keywords, we collected tweets with metadata for seed words - *#stateofwomen*, *#itsonus*, and *#heforshe*, for 7 days after the launch of *#StateOfWomen*, on Jun 14 2016. We extract and store all the relevant metadata such as tweet text, timestamp, and author profile information such as full name, and location. Table 1 provides the basic dataset statistics.

Campaign Community	# Tweets	Retweets (%)	# Authors
All	168,950	124,952 (74%)	72,957
#ItsOnUs	1,415	1,048 (74%)	1,095
#StateOfWomen	157,288	115,467 (73%)	66,615
#HeForShe	9,112	7,564 (83%)	6,452
#ItsOnUs & #StateOfWomen	890	734 (82%)	656
#ItsOnUs & #HeForShe	5	1 (20%)	5
#StateOfWomen & #HeForShe	239	138 (58%)	97
#ItsOnUs & #StateOfWomen & #HeForShe	1	0 (0%)	1

Table 1: Sets of various combinations of campaign related tweets.

User Type: Individual vs. Organization. Researchers in computational social science have investigated various methods to classify user types on Twitter, such as those based on demographic attributes, influence, and ideologies [14]. From an organizational study perspective, we choose to explore user identity type of *organization* versus *individual*. An organizational type of user account is one that represents a group, company or an organization, such as a NGO account @EndRapeOnCampus on Twitter. All the other users are considered individual accounts. We adapted the approach of [13] for individual versus organization user classification that provided an implementation on GitHub. It takes input to classify a user as the set of tweet metadata objects, including tweet and the user information as returned by the Twitter API. Manual verification of randomly selected 50 users gave an accuracy of 84%. We expect that organizational user types are likely to be fewer in number than individuals in general given their numbers in real-world and table 2 reflects such a distribution.

	# Users		# Tweets	
Total	72,957		168,950	
Organization	86	0.11%	8,574	5%
Individual	72,871	99.88%	160,376	94.9%

Table 2: User Type Distribution.

User Gender: Male vs. female vs. Unisex. For the gender classification process, we only consider users that were classified as individuals in the preceding step. Twitter does not ask users to provide gender in the registration process, and therefore, several studies have proposed methods to infer a user’s gender on Twitter [14]. The existing implementation from prior studies using content and profile features was not found to be efficient in our manual verification study of randomly selected 50 users, such as maximum accuracy of 46% when using prior work approach [14]. Therefore, we resolved to use name-based gender identification, at the expense of poor dataset coverage. We used user-profile metadata for user’s real-name, whenever available. This tool is available on GitHub and provides reference to the name database for each country and also the database of first names from all around the world provided together with an open source C program for name-based gender inference (<https://goo.gl/B7QLJx>). We found accuracy of 78% in the manual verification. Table 3 shows statistics by gender

and we note the consistency with literature on the higher female participation for anti-GBV activism in both online and offline environments.

User Location. We use the author profile location field for a user to study the geographical engagement in the campaigns across US region. We map profile location metadata to the US states by querying the textual location field values in open source search tool Nominatim (<https://goo.gl/wbwfM0>) based on OpenStreetMap data. For the 47% of users, we resolved the location metadata given the noisy metadata (e.g., ‘IN THE FREE MARKET’). We use user locations by states to study the patterns of state level campaign engagement.

	# Individuals	Tweets		
Total	72,871		160,376	
Male	13,826	18.97%	25,045	15.6%
Female	30,733	42.2%	69,289	43.2%
Unisex	1,302	1.8%	2,895	1.8%
Unknown	27,010	37.1%	63,147	39.4%

Table 3: Gender-wise distribution of Individual Users.

Research Question Analysis

R1: Crossover of Campaign Communities. For finding overlapping engagement by the users in the different campaign communities, we first computed sets of users engaging in campaigns and then, Jaccard Similarity coefficient (table 4) across campaign pairs, which measures similarity between finite sample sets.

Community A	Community B	All Users	Organizational Users
#ItsOnUS	#StateOfWomen	0.02	0.49
#ItsOnUS	#HeForShe	0.01	0.15
#StateOfWomen	#HeForShe	0.01	0.13

Table 4: Campaign Community Overlap of users via Jaccard Coefficient.

Results in Table 4 show a weak overlap between the members of the campaign communities, especially with the global campaign *#HeForShe*, despite having related social causes. Engaged organizations appear to be more aware and connected in the case of college and nation-centric campaigns, partly due to their potentially common supporting audience (both are White House-launched initiatives). For studying the cross-engagement of users by cross-referencing campaign identity hashtags in tweets, we created 7 subsets of data as the following: Tweets containing a.) only *#ItsOnUs*, b.) only *#StateOfWomen*, c.) only *#HeForShe*, d.) both *#ItsOnUs* and *#StateOfWomen*, e.) both *#ItsOnUs* and *#HeForShe*, f.) both *#StateOfWomen* and *#HeForShe*, and g.) tweets containing all the three *#ItsOnUs*, *#StateOfWomen* & *#HeForShe*. Table 1 shows the results for the tweet volumes under these 7 sets. We found only one user to be intersecting all the three campaigns. Along the same line, although the cross-referencing for college and nation-centric campaigns *#ItsOnUs* and *#StateOfWomen* exists (albeit less than 10% compared to individual campaign tweet volumes), there is a lack of much intersection with the global campaign.

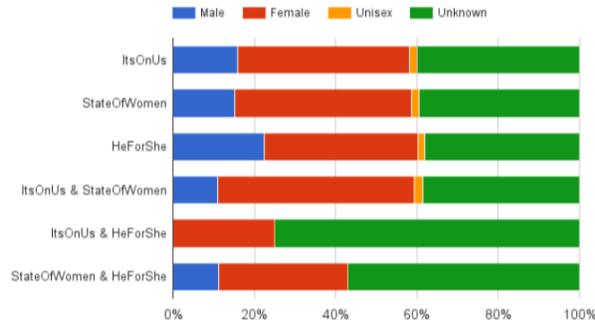


Fig. 1: Gender-wise relative distribution of tweets across campaign communities.

R2: Demographic Effect on Campaign Engagement. Figure 1 shows the relative distribution of generated tweets across campaign communities grouped by the author’s gender. We noticed that the college-centric campaign #ItsOnUs has a similar pattern with #StateOfWomen, while not as much with #HeForShe. Although there is a consistent pattern of higher female engagement across all three campaigns, we can note the higher proportions of male engagement in #HeForShe. This should motivate campaigns to coordinate with other campaigns, as coordination could improve strategies in engaging various demographics to generate awareness and content.

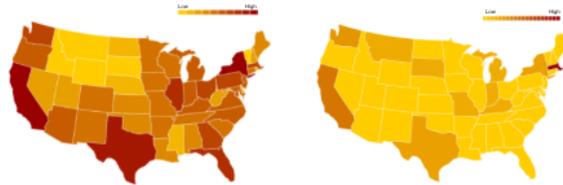


Fig. 2: Max-normalized volume of users in the anti-GBV campaigns (left) and the 2014 GBV-related crime reports from FBI UCR (right).

We also computed correlation between the state-wise statistics across US for GBV-related crimes data from the FBI Uniform Crime Reporting database (reported hate-crime statistics in 2014 by selecting categories of *Gender* and *Gender Identity*, available at <https://ucr.fbi.gov/>) and the volume of participating users in campaigns from those states, to measure the regional engagement patterns (Figure 2). We noticed a positive correlation coefficient of 0.34, indicating anti-GBV campaign engagement and awareness in consistence with the regional crime rates. There is an opposite trend for few states (e.g., South Dakota), indicating the need for region-aware strategies for the awareness campaigns. We also found higher female participation than the male participation across states and across campaigns, indicating a greater need for engaging male and unisex users.

Conclusion & Future Direction

We presented a preliminary study of user engagement in anti-GBV campaign communities on social media, by analyzing three diverse campaigns on Twitter

via cross-community participation of user types and gender as well as a comparison of regional crime rates with Twitter engagement in the campaigns. This study will help inform more intensive domain-driven content and interaction network analysis of interrelated, cause-driven campaigns and coordination of organizational users in them. Likewise, in the offline movement of violence against women that crystallized in the 20th century where activist partners had "learned a strategic lesson - *Let's look for more allies*", in part due to the desire to link the call for women's rights to the wider issue of human rights overall [15]. This preliminary analysis examined and relied on existing user metadata inference methods, such as gender classifier to analyze the dataset. Improving the accuracy of these classifiers and detecting errors in self-reported data provides a direction for future work. In future, we will also investigate linguistic factors that motivate deeper engagement beyond information sharing, as suggested in Table 1, tweets including both *#StateOfWomen* and *#HeForShe* had lesser retweets (58%) but more original posts than other cases.

References

1. Gender Bases Violence. United Nations - UNWomen.Org. HeForShe Retrieved from <http://goo.gl/QZ2baP>. (2016)
2. Russo, Nancy, F., Pirlott, A.: Gender-based violence. *Annals of the new york academy of sciences*, Vol. 1087(1),178-205 (2006).
3. It's On Us Campaign. ItsOnUs Retrieved from <http://goo.gl/xponHJ>. (2016).
4. The United States Of Women Campaign. StateOfWomen Retrieved from <http://goo.gl/JsxwpL> (2016).
5. Centers for Disease Control and Prevention. National Intimate Partner Sexual Violence Survey: 2010 Summary Report. Available at <http://goo.gl/DxTAqW> (2010).
6. American Association of Universities. Campus Climate Survey on Sexual Assault and Sexual Misconduct. Retrieved from <https://goo.gl/ABuvzZ> (2015).
7. De Choudhury, M., Jhaver, S., Sugar, B., Weber, I. Social Media Participation in an Activist Movement for Racial Equality. In ICWSM-16 (2016).
8. Conover, M.D., Ferrara, E., Menczer, F., Flammini, A. The digital evolution of occupy wall street. *PloS one*, Vol. 8(5), e64679. (2013).
9. Hewitt, S., Tiropanis, T., Bokhove, C. The problem of identifying misogynist language on Twitter and other online social spaces. In *Web Science*, 333-335 (2016).
10. Purohit, H., Banerjee, T., Hampton, A., Shalin, V. L., Bhandutia, N., Sheth, A. Gender-based violence in 140 characters or fewer: A *#BigData* case study of Twitter. *First Monday*, 21(1). (2016).
11. Messner, Michael A., Max A. Greenberg, T.P. *Some men: Feminist allies and the movement to end violence against women*. Oxford University Press. (2015).
12. Lovejoy, K., Waters, R. D., Saxton, G. D. Engaging stakeholders through Twitter: How nonprofit organizations are getting more out of 140 characters or less. *Public Relations Review*, Vol. 38(2),313-318 (2012).
13. McCorriston, J., Jurgens, D., Ruths, D. Organizations Are Users Too: Characterizing and Detecting the Presence of Organizations on Twitter. In ICWSM. (2015).
14. Volkova, S., Coppersmith, G., Van Durme, B. Inferring User Political Preferences from Streaming Communications. In *ACL* (1) (pp. 186-196) (2014).
15. Keck, Margaret E. and Kathryn Sikkink. *Activists Beyond Borders: Advocacy Networks in International Politics*. Ithaca: Cornell University Press. (1998).