2022 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)

The Pursuit of Being Heard: An Unsupervised Approach to Narrative Detection in Online Protest

Kumari Neha*, Vibhu Agrawal*, Arun Balaji Buduru*, Ponnurangam Kumaraguru†
*Indraprastha Institute of Information Technology, Delhi, India
†International Institute of Information Technology, Hyderabad, India
{nehak, vibhu18116, arunb}@iiitd.ac.in
pk.guru@iiit.ac.in

***Abstract*—Protests and mass mobilization are scarce; however, they may lead to dramatic outcomes when they occur. Social media such as Twitter has become a center point for the organization and development of online protests worldwide. It becomes crucial to decipher various narratives shared during an online protest to understand people's perceptions. In this work, we propose an unsupervised clustering-based framework to understand the narratives present in a given online protest. Through a comparative analysis of tweet clusters in 3 protests around government policy bills, we contribute novel insights about narratives shared during an online protest. Across case studies of government policy-induced online protests in India and the United Kingdom, we found familiar mass mobilization narratives across protests. We found reports of on-ground activities and call-to-action for people's participation narrative clusters in all three protests under study. We also found protest-centric narratives in different protests, such as skepticism around the topic. The results from our analysis can be used to understand and compare people's perceptions of future mass mobilizations.**

***Index Terms*—Social Media Protest, Unsupervised clustering, Protests, Narratives, Twitter**

I. Introduction

Social media has become integral to various social movements, and protests due to easy information dissemination and wider public reach [1]–[5]. Irrespective of the different socio-economic circumstances or political agendas, the various online protests share similar morphological features in using social media for self-organization and obtaining a more significant number of participants [6]. Using a hashtag to build a collective narrative makes Twitter one of the prime spots for conducting protest [7]. While Twitter enables a broad reach of the protest, a fine-grained analysis of various narratives present within a protest setting may also help decipher the people's perception and shed light on people's will and social protest's overall focus.

Previous studies on social media movements/protests have focused on different collective narratives in the campaign [4], [8]. The narratives range from information dissemination (such as personal grievances) around the topic [3], [9]; to call for participation [10] or reporting of on-ground activities [11], as shown in Figure 1. The

Fig. 1: Figure showing examples of different narratives expressed by people during online protests. CTA: Call-to-action, OGA: On-ground activities, GRV: personal grievances.

grievance narrative might include personal stories of perceived injustice or other forms of hardships related to the cause. On-ground activities are narrative that either comes from people who are witnessing the offline protest or posts about current online activities related to the protest. The call for participation (call-to-action) narrative urges the users to participate in the cause by either being part of the physical protest or using social media to tweet protest-related posts. Although the different narratives during a protest have been studied individually, a unified discussion of various narratives present within a protest is scarce [4].

In this paper, we focus on various narratives in recent instances of the Reform movement [12] in India and the UK, where policies introduced by the government in power were deemed unjust and demanded to be repealed [13]–[15]. The reform movements studied in this work are as follows -

Citizenship Amendment Act, 2019 (CAA): The Citizenship Amendment Act, 2019 was passed by the Indian Government on December 11, 2019. It allows illegal immigrants who have faced religious persecution in Afghanistan, Bangladesh, or Pakistan to seek citizenship in India if they have entered India on or before December 31, 2014 [16]. This led to a protest in the country with a debate on the non-secular roots of the Act.

IEEE/ACM ASONAM 2022, November 10-13, 2022
978-1-6654-5661-6/22/$31.00 É 2022 IEEE

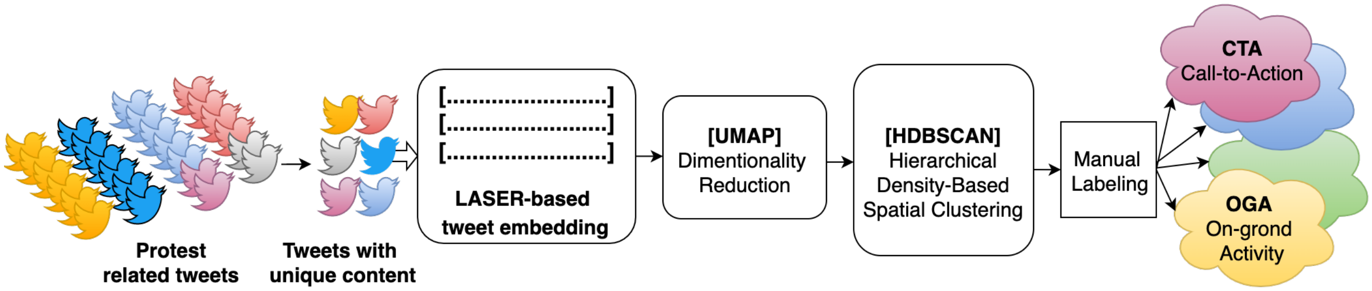


Fig. 2: Framework to identify dominant narratives amid social media protest. The different color of tweet represents different narrative tweets present in the dataset.

Farmer's Protest, 2020 (FP): The Indian government proposed the Farmer's bill on September 20, 2020. The country's farmers feared that the three laws introduced in the bill would abolish the minimum support price (MSP), leaving farmers at the mercy of big corporations. The turn of events in the country led the Indian government to finally repeal the bill on November 09, 2021, ending the year-long protest in the country [13].

Kill the Bill Protest, 2022 (KTB): The Police, Crime, Sentencing, and Courts Bill (PCSC) introduced new police powers and reviewed the present rules around crime and protests in England and Wales. The activists opposed the law due to its ability to impose conditions on any protest deemed disruptive to the local community [15].

The primary motivation for using cluster-based analysis is to leverage the semantic difference between clustered texts and identify fine-grained separation between clusters as different narratives in a protest. Using a clustering-based framework, we bridge the gap of unified narrative detection in social media protests and identify converging narratives across different protests. Broadly, we ask the following research questions:

RQ 1: What are the different narratives present in a protest?

RQ 2: What are the most prominent narratives present within a protest?

RQ3: Are there any converging narratives across protests?

The succeeding sections of the paper are organized as follows. We discuss the Data and Methods in Section II, followed by Results in Section III and the Conclusion in Section IV.

II. Data and Method

This section discusses the dataset and method for discovering the narrative clusters in the protest tweets.

A. Data

For CAA, we use the data collected in [17] for our study. For FP and KTB, we used the trending hashtags during the debate to collect protest-relevant tweets.

The statistics of the collected data are present in Table I. For initial tweet preprocessing, we used the methods presented in [17]. After the first pre-processing step, the CAA tweet count was $11,302,023$, for FP it was $1,500,022$ and for KTB it was $278,065$. We release the anonymized version of our data at https://precog.iiit.ac.in/resources.html.

TABLE I: Statistics of the data used to analyze campaign narratives. Abbr; CAA: Citizenship Amendment Act, FP: Farmers Protest, KTB: Kill the Bill protest.

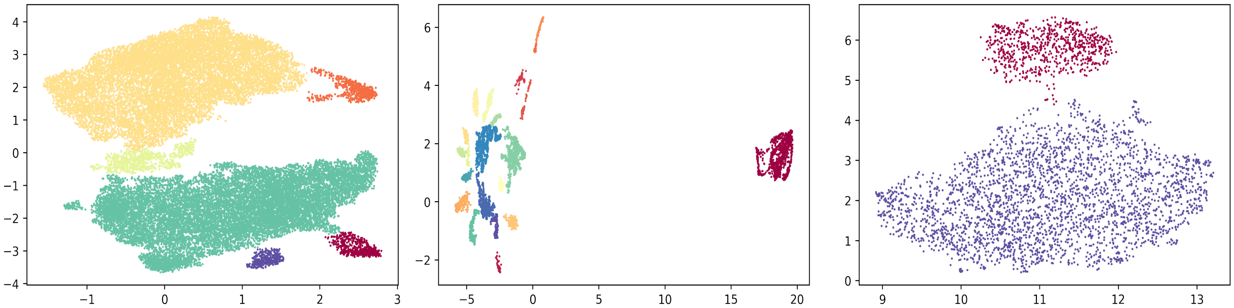
Protest	Start date	End date	#Tweets	#Users
CAA	Dec 07, 2019	Feb 27, 2020	$11,350,276$	$931,175$
FP	Mar 31, 2021	Aug 13, 2021	$1,509,703$	$160,286$
KTB	Jan 14, 2022	Jan 26, 2022	$280,549$	$73,666$

B. Method:

The narratives in an online campaign tend to be topic-driven [8], [18]; therefore, a fixed set of labels may not always fit a given protest. Hence, we propose an unsupervised framework for identifying significant narratives of a protest, as shown in Figure 2. Our framework is inspired by the unsupervised user-based stance detection framework proposed recently [19], [20].

Active tweet identification: We use a two-step process to consider a rich and unique instance of tweets around a protest. First, we use string matching to identify duplicate tweets in an online protest. We remove the hashtags and mentions to conduct string matching on the tweet text. The practice of tweeting the same text instead of retweeting has recently gained much traction in the global south recently [21]. Secondly, we use tweets whose occurrence (duplicates) exceeds a particular threshold based on the size of the data, and manual intervention, where we recheck the cluster outputs with different threshold values.

Tweet representation: To identify the most active tweets for CAA, we chose the threshold for semantic similarity as 30 and found $36,109$ tweets that account for $7,878,996$ tweets/retweets. For FP, we selected the threshold as 30 and obtained $7,553$ unique tweets that belonged to $112,186$ total tweet/retweet. The threshold



(a) #CAA (b) #FP (c) #KTB

Fig. 3: Clusters of narratives for CAA, FP, and KTB, respectively.

for KTB was set to 5 and resulted in $3,821$ tweets that account for a total of $278,065$ tweets. Once we have identified the most active tweets, we first represent the tweets in the embedding space using BiLSTM encoder-based universal language agnostic sentence embedding (LASER) [22], which has proven to give the best performance for retaining linguistic information among various sentence embeddings [23].

Tweet projection: We then project each tweet onto a two-dimensional plane using Uniform Manifold Approximation and Projection (UMAP) algorithm [24]. UMAP attempts to project similar elements closer to each other while dissimilar elements are projected far away.

Clustering: We cluster the projected tweet vectors using hierarchical density-based clustering (HDBSCAN) [25]. HDBSCAN finds clusters of varying densities. Among clustering algorithms, including Meanshift [26] and DBSCAN [27], HDBSCAN gave us the best clustering performance, determined by manual evaluation. We used prominence score to analyze and validate our manual labeling of different narratives qualitatively.

III. Results

RQ1: Narratives present in a protest

Per RQ 1, we examine the clusters formed in each campaign using our framework. We have not reported the tweets clustered as noise for brevity. To annotate protest clusters into different narratives, we leverage the previous literature on protest studies in different parts of the world [5], [9], [10].

CAA: As shown in Figure 3a, 6 clusters of tweets were formed for CAA. To analyze narratives, we manually annotate randomly selected two sets of 10 sample tweets from each cluster. Table II shows the 4 different narratives clusters in the campaign with highest engagement. The other two clusters belonged to personal grievances and location-specific tweets. In terms of engagement (i.e., tweet/retweet activity), the largest cluster showed skepticism towards the Act (e.g., Thousands on the street in support of CAA! I was not expecting this from Bhubaneshwar). The second dominant narrative for CAA was the Questioning cluster, where the tweets posed questions to the Act, politicians, and protesters for violent actions. The other two important narrative clusters included call-to-action and on-ground activities clusters.

FP: We found 20 clusters for FP. However, we focused on the top 4 clusters for further analysis, constituting more than 500 unique tweet text each. As shown in Table II, the most dominant narrative in FP was call-to-action, with $6,287$ (CTA-AP) and 845 (CTA-AP) unique tweets respectively. While the cluster (denoted as CTA) called for participation in support of farmers, the cluster CTA-AP (i.e., Call to action against politicians) contained tweets against the ruling government for their proposal of the bill.

KTB: The UK protest on the policing formed 2 clusters using our framework as shown in Figure 3c. Among the two clusters, more engagement was around call-to-action.

RQ2: Prominent narratives in a protest

We found the presence of call-to-action (CTA) and reporting of on-ground activities (OGA) forming two persistent clusters in all the protests under study. The other common narrative across protests is grievances or personal complaints [9]. Our proposed framework was able to form clusters with deductible characteristics for call-to-action and on-ground reporting of activities with similar features across the protests under study. The skepticism and questioning in CAA reveal the contention in the online social media about the Act. On the contrary, the FP and KTB protests were more in harmony with opposing the bill, with narratives formed majorly towards CTA and OGA.

RQ3: Converging narratives across protests

The converging narratives across protests were call-to-action (CTA) and on-ground activities (OGA). With the help of Prominence score Pr in Equation 1, we found the most prominent terms, emojis, hashtags, and mentions in each narrative cluster. The Prominence score also helped validate the cluster narratives identified through manual

TABLE II: Main narratives present in the protests under study.

Protest	Narrative	Unique Tweets	#Tweets	#Users
	Question-ing	13,380	2,387,533	278,184
CAA	Skepticism	15,274	3,911,679	466,139
	CTA	865	154,926	72,415
	OGA	647	98,221	48,276
	CTA	6,287	13,734	464
FP	CTA-AP	845	26,897	9,470
	OGA	683	66,660	2,538
	OGA	742	20,557	9,431
KTB	CTA	2,958	178,499	56,079

annotations. Table II shows the engagement around the prominent narratives in the protests under study.

A. Qualitative analysis of clusters:

This section sheds light on the framework's performance. We focus on the semantic difference between clusters as a measure of cluster quality [19], [20]. We identify the most prominent term, emojis, hashtags, and mentions in each cluster to show how each narrative uniquely talks about the same issue in a different context. To suit our need, we generalize the prominence score used in the literature [20] for more than two cases.

For each term t, we capture the degree of its occurrence in a set of tweets from cluster i, i.e., tf_i, as compared to all other clusters tf_j (where j ranges from $Cluster_1$ to $Cluster_n$). The *prominence score* of a term t is defined as a product of its valence score and its term frequency as follows:

$$V(t,i) = \log(tf_{t,i}) * (2 * \frac{\frac{tf_i}{total_i}}{\sum_{j=1}^{n} \frac{tf_j}{total_j}} - 1) \tag{1}$$

We discuss the result for each protest in detail below: **CAA:** For OGA, the top terms include state and location information. It gives evidence of users sharing location-specific on-ground activity on social media. The top hashtags include states in India (i.e., Assam, Uttar Pradesh, etc.). The top terms and hashtags show the prominence of Indian states mentioned in the cluster. The top emojis used in OGA was 👈, 📢, 🚩, and 🖤. The top mention in OGA included news personnel and ministers. For CTA, the top terms include words like 'initiative', 'showcase', 'trending', and 'notion of 'we''. The top hashtags include a pledge to solidarity (#solidaritypledge). Most of the top accounts under CTA are currently suspended by Twitter. At the same time, others included political party leaders. In CAA, we found that the OGA narrative more prominently mentioned news channel personnel, while common people were mentioned mostly in CTA.

FP: The top terms for the OGA narrative for FP include 'arrest', 'missing', and locations of on-ground activity. The most prominent emojis include 🖤, 💔, 🎥, and 💵. The OGA narrative's prominent mention included NGO handles, politicians, and news outlets. CTA's top terms and hashtags included appreciation and farmer's pride. CTA-AP included terms like 'nazi' and 'socialism'. The context-specific emojis of crops (🌾), farmers (👩‍🌾), and tractors (🚜) were commonly used in CTA and CTA-AP. Prominent mentions in CTA-AP were of Bollywood actors, farmer's unions, and activists. CTA mentioned the prime minister among other activist accounts and a few suspended accounts.

KTB: The top prominent terms and hashtags for OGA in the KTB protest included narratives of shame against the Prime Minister and reporting of deaths. The emojis used include 😤,🔥,🤦‍♀️. The mentions in OGA included the Prime minister's handle, members of parliament, and other politicians. While the CTA cluster's top terms and hashtags included words like peers, places, and calling out activists, the top mentions included members of the green party and activists' handles. The top emojis include ‼️, ⚠️, 🖤, 👀, 🎥.

IV. Conclusion

This work proposes an unsupervised framework to identify the different narrative clusters in a social media protest. Catering to the fact that protests are composed of certain narratives discussed in previous literature, we leverage clustering algorithms to cluster protest narratives. We used the anti-government policy bill-related tweets in India and the United Kingdom and deciphered the most prominent and common narratives within and across the protests. The proposed Prominence score validation for narratives is qualitatively consistent in all protests under study. We found that call-to-action and on-ground activities as converging narratives across protests. In a protest that led to discourse, we found narratives that show skepticism and questioning tweets. However, we can conclude that the protests that contain majorly on-ground reporting and call-to-action are single-sided anti-government protests. With the help of the prominence score, we found a pattern of emojis commonly used in protest-related tweets. The mentions in the protests provide evidence that OGA has more verified accounts tagged, while the CTA mentions more of the general public, some of whom were suspended.

References

[1] R. Korolov, D. Lu, J. Wang, G. Zhou, C. Bonial, C. Voss, L. Kaplan, W. Wallace, J. Han, and H. Ji, "On predicting social unrest using social media," in *2016 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM)*. IEEE, 2016, pp. 89–95.

[2] M. De Choudhury, S. Jhaver, B. Sugar, and I. Weber, "Social media participation in an activist movement for racial equality," in *Proceedings of the... International AAAI Conference on Weblogs and Social Media. International AAAI Conference on Weblogs and Social Media*, vol. 2016. NIH Public Access, 2016, p. 92.

[3] A. Field, G. Bhat, and Y. Tsvetkov, "Contextual affective analysis: A case study of people portrayals in online #metoo stories," *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 13, pp. 158–169, Jul. 2019. [Online]. Available: https://www.aaai.org/ojs/index.php/ICWSM/article/view/3358

[4] R. Wang and A. Zhou, "Hashtag activism and connective action: A case study of# hongkongpolicebrutality," *Telematics and Informatics*, vol. 61, p. 101600, 2021.

[5] G. Lotan, E. Graeff, M. Ananny, D. Gaffney, I. Pearce *et al.*, "The Arab Spring| the revolutions were tweeted: Information flows during the 2011 Tunisian and Egyptian revolutions," *International journal of communication*, vol. 5, p. 31, 2011.

[6] S. González-Bailón, J. Borge-Holthoefer, A. Rivero, and Y. Moreno, "The dynamics of protest recruitment through an online network," *Scientific Reports*, vol. 1, pp. 1–7, 2011.

[7] R. Wang and K.-H. Chu, "Networked publics and the organizing of collective action on Twitter: Examining the #Freebassel campaign," *Convergence*, vol. 25, no. 3, pp. 393–408, 2019.

[8] K. Neha, T. Mohan, A. B. Buduru, and P. Kumaraguru, "Truth and travesty intertwined: a case study of# ssr counterpublic campaign," in *Proceedings of the 2021 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 2021, pp. 643–648.

[9] A. Sinpeng, "Hashtag activism: social media and the #freeyouth protests in thailand," *Critical Asian Studies*, vol. 53, no. 2, pp. 192–205, 2021. [Online]. Available: https://doi.org/10.1080/14672715.2021.1882866

[10] A. Rogers, O. Kovaleva, and A. Rumshisky, "Calls to action on social media: Potential for censorship and social impact," *EMNLP-IJCNLP 2019*, p. 36, 2019.

[11] O. Varol, E. Ferrara, C. L. Ogan, F. Menczer, and A. Flammini, "Evolution of online user behavior during a social upheaval," in *Proceedings of the 2014 ACM conference on Web science*, 2014, pp. 81–90.

[12] J. DeFronzo and J. Gill, *Social problems and social movements*. Rowman & Littlefield, 2019.

[13] E. W. desk, "Farmers end year-long protest: A timeline of how it unfolded," 2021. [Online]. Available: https://indianexpress.com/article/india/one-year-of-farm-laws-timeline-7511961/

[14] V. S. Damini Nath, "After a heated debate, Rajya Sabha clears Citizenship (Amendment) Bill," *The Hindu*, 2019.

[15] T. B. I. W. desk, "What are the Kill the Bill protests?" *The Big Issue*, 2022.

[16] A. Chandrachud, "Secularism and the citizenship amendment act," *Indian Law Review*, vol. 4, no. 2, pp. 138–162, 2020. [Online]. Available: https://doi.org/10.1080/24730580.2020.1757927

[17] K. Neha, V. Agrawal, V. Kumar, T. Mohan, A. Chopra, A. B. Buduru, R. Sharma, and P. Kumaraguru, "A tale of two sides: Study of protesters and counter-protesters on #citizenshipamendmentact campaign on twitter," in *14th ACM Web Science Conference 2022*, ser. WebSci '22. New York, NY, USA: Association for Computing Machinery, 2022, p. 279–289. [Online]. Available: https://doi.org/10.1145/3501247.3531584

[18] A. Panda, R. Kommiya Mothilal, M. Choudhury, K. Bali, and J. Pal, "Topical focus of political campaigns and its impact: Findings from politicians' hashtag use during the 2019 indian elections," *Proceedings of the ACM on Human-Computer Interaction*, vol. 4, no. CSCW1, pp. 1–14, 2020.

[19] K. Darwish, P. Stefanov, M. J. Aupetit, and P. Nakov, "Unsupervised User Stance Detection on Twitter," 2019. [Online]. Available: http://arxiv.org/abs/1904.02000

[20] A. Rashed, M. Kutlu, K. Darwish, T. Elsayed, and C. Bayrak, "Embeddings-Based Clustering for Target Specific Stances: The Case of a Polarized Turkey," 5 2020. [Online]. Available: http://arxiv.org/abs/2005.09649

[21] M. Jakesch, K. Garimella, D. Eckles, and M. Naaman, "Trend alert: A cross-platform organization manipulated twitter trends in the indian general election," *Proc. ACM Hum.-Comput. Interact.*, vol. 5, no. CSCW2, oct 2021. [Online]. Available: https://doi.org/10.1145/3479523

[22] M. Artetxe and H. Schwenk, "Massively multilingual sentence embeddings for zero-shot cross-lingual transfer and beyond," *Transactions of the Association for Computational Linguistics*, vol. 7, pp. 597–610, 2019.

[23] K. Krasnowska-Kieraś and A. Wróblewska, "Empirical linguistic study of sentence embeddings," in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 2019, pp. 5729–5739.

[24] L. McInnes, J. Healy, and J. Melville, "Umap: Uniform manifold approximation and projection for dimension reduction," *arXiv preprint arXiv:1802.03426*, 2018.

[25] L. McInnes and J. Healy, "Accelerated hierarchical density based clustering," in *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*. IEEE, 2017, pp. 33–42.

[26] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 5, pp. 603–619, 2002.

[27] E. Schubert, J. Sander, M. Ester, H. P. Kriegel, and X. Xu, "Dbscan revisited, revisited: why and how you should (still) use dbscan," *ACM Transactions on Database Systems (TODS)*, vol. 42, no. 3, pp. 1–21, 2017.