Topical Focus of Political Campaigns and its Impact: Findings from Politicians' Hashtag Use during the 2019 Indian Elections

ANMOL PANDA, Microsoft Research India
RAMARAVIND KOMMIYA MOTHILAL, Microsoft Research India
MONOJIT CHOUDHURY, Microsoft Research India
KALIKA BALI, Microsoft Research India
JOYOJEET PAL, Microsoft Research India

We studied the topical preferences of social media campaigns of India's two main political parties by examining the tweets of 7382 politicians during the key phase of campaigning between Jan - May of 2019 in the run up to the 2019 general election. First, we compare the use of self-promotion and opponent attack, and their respective success online by categorizing 1208 most commonly used hashtags accordingly into the two categories. Second, we classify the tweets applying a qualitative typology to hashtags on the subjects of nationalism, corruption, religion and development. We find that the ruling BJP tended to promote itself over attacking the opposition whereas the main challenger INC was more likely to attack than promote itself. Moreover, while the INC gets more retweets on average, the BJP dominates Twitter's trends by flooding the online space with large numbers of tweets. We consider the implications of our findings hold for political communication strategies in democracies across the world.

CCS Concepts: • **Human-centered computing** → **Social media**; *Collaborative content creation*; *Computer supported cooperative work*; *Social networking sites.*

Additional Key Words and Phrases: Twitter; India; Politics; Hashtags; Election Campaigns; Narendra Modi; Rahul Gandhi; Political Communication; Polarization

ACM Reference Format:
Anmol Panda, Ramaravind Kommiya Mothilal, Monojit Choudhury, Kalika Bali, and Joyojeet Pal. 2020. Topical Focus of Political Campaigns and its Impact: Findings from Politicians' Hashtag Use during the 2019 Indian Elections. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW1, Article 53 (May 2020), 14 pages. https://doi.org/10.1145/3392860

1 INTRODUCTION

Following the lead of Barack Obama's historic 2008 election victory in the US, politicians across the world have leveraged social media to reach out to the electorate. In recent years, this trend has extended to bypassing the filters of the mainstream press. Heads of state including Donald Trump and Jair Bolsonaro have moved to communicating primarily on social media, partly by demonizing the mainstream media as biased, instead referring to social media as a reflection of

Authors' addresses: Anmol Panda, anmol.panda777@gmail.com, Microsoft Research India, Vigyan #9, Lavelle Road, Bengaluru, Karnataka, 560 001; Ramaravind Kommiya Mothilal, t-rakom@microsoft.com, Microsoft Research India; Monojit Choudhury, Microsoft Research India, monojitc@microsoft.com; Kalika Bali, Microsoft Research India, kalikab@microsoft.com; Joyojeet Pal, Microsoft Research India, joyojeet.pal@microsoft.com.

2573-0142/2020/5-ART53 $15.00
https://doi.org/10.1145/3392860

true public sentiment. In India, the phenomenon of social media use by politicians had a major boost following the landslide victory of Narendra Modi's Bharatiya Janata Party (BJP) in the 2014 elections, a campaign which had a critical online component.

India is a parliamentary democracy, and has since its independence in 1947, primarily had one dominant political party - the Indian National Congress (INC). However, the INC has slowly declined in its parliamentary seat share, and since the 1984 general elections, no party secured an outright majority, relying instead on pre-poll agreements and coalitions for government formation. This was changed with the BJP's 2014 victory, the biggest single party performance in three decades. Much work has argued that the 2014 campaign was highly personality-centric [11], and that campaigns were fundamentally metamorphosized in India by Modi's aggressive use of social media [13], marked in particular by his exclusive use of direc online missives, primarily on Twitter, to communicate with the professional journalist corps [36].

The 2014 election was also important for its aftermath, which saw new technology-mediated campaigns turning into the norm. While the parliamentary system had meant that parties traditionally campaigned on party symbols and ideology rather than individual leaders, the success of Modi created an industry of social media campaign specialists joining politicians and parties[28]. The move towards social media enabled elections was strengthened by the litmus test of the 2017 Uttar Pradesh elections, which saw massive use of social media propaganda and online misinformation [32]. Uttar Pradesh, the country's largest state, is also one of the poorest and most backward, and the widespread use of social media, in particular WhatsApp, underlined parties' acceptance that this was no longer an elite phenomenon and that it had become a central weapon of election campaigns. In Modi's own pre-2019 elections diktat to his party, he required that anyone looking for a ticket to contest elections for the BJP needed to show a minimum threshold of followers online [47, 51]

A consequence of these changes has also been much discussion over whether the political speech has itself moved towards more polarized rhetoric with unmediated online debates increasingly central to the overall shaping of political communication [33]. Our work seeks to systematically examine digital outreach across parties, through a large-scale snapshot of the topics and scope of online political speech in the 2019 elections.

We built a database of over 18500 Indian politicians on Twitter, and studied a selection of subject matters addressed in their output during the 2019 general elections. We based our analysis on the choice of hashtags posted by politicians of India's ruling Bharatiya Janta Party (BJP) and main opposition Indian National Congress (INC) party. We restricted our study to tweets containing hashtags due to the critical affordances they presented to the author: networking beyond their followers [6, 14] and participating in a collective effort to promote a cause [10].

Our findings indicate marked differences between the issues each party focused on and their success in securing traction on Twitter for their narrative.

In section 2, we discuss prior work that used hashtags to gauge partisanship and model tweet topics. In section 3, we describe the dataset and the hashtag typology we have used. Section 4 lists the results of our analysis. In section 5, we probe the implications of our study for political communication on social media. Finally, we discuss the limitations and potential future directions of work using this dataset.

2 RELATED WORK

There is a large body of literature on Twitter political campaigns [9, 20, 22, 25, 29, 35, 49, 50]. Hemphill et al. presented a framework to study partisanship of political campaigns on Twitter through hashtags [18, 19] and used it to study political framing [18] and partisan messaging in US Congressional elections [3, 21]. Hashtags have been used to analyse Twitter activity of political elite in Sweden [30] and links between politicians' Twitter messaging and their coverage by journalists on

Twitter in Norway [14] and the mainstream media in the US [46]. Lunde [31] used hashtags to study political humour on Russian Twitter while Mirko et al. analysed polarization [27] among followers of members of US congress and Governors. Researchers have used thematic typologies to study insults [38], negative messaging [8], and topical preferences [1, 5, 53] in political communication.

Prior literature has considered hashtag use by national leaders, members of parliament and provincial Governors. We build upon these works by conducting the first large scale study of Indian politicians' hashtag usage during a national election campaign. The scale of our study is much larger than what prior studies have attempted and fills a void that has hitherto escaped researchers attention.

Within the CSCW community, researchers have studied online political communication at length. Grevet et al. [16] looked at the correlation between political homophilly and weak ties on Facebook and suggest measures to connect politically distant users to reduce polarization. Borge-Holthoefer[4] et al. studied polarization on Twitter in Egyptian political discourse whereas Semaan[45] et al. recommended measures to reduce the same using a qualitative study of 21 US citizens. Furthermore, Kulshreshtha[26] et al. have quantified political bias in searches on Twitter while Park[40] et al. attempted to predict orientation of news stories based on sentiment patterns of comments posted by viewers. In addition, existing literature has considered collaborative political blogging[2], political activists' organization practices online[44] and the use of storytelling in social movements on new media[12].

Our work dovetails into this extant body of work in that it focuses not on the tweet patterns of individual political actors, but the collective output of collaborative propaganda efforts by members of political organizations. In large and diverse political systems like India, creating and enforcing message discipline in national election campaigns is a difficult feat. However, social media platforms have afforded organizations new ways of collaborative action. The size and diversity of our sample, therefore, provides a unique snapshot of nationwide campaigns, divided by a diverse, multi-lingual polity but unified in their messaging towards the shared goal of winning elections.

3 DATA

We defined a politician on Twitter as a public figure holding a position within a political party. This included elected members of the union parliament, state legislatures and local governing bodies. We also included unelected party officials like national, state and district party presidents and vice-presidents, spokespersons, general secretaries. Lastly, we included youth-wing (IYC for the INC and BJYM for the BJP) and student-wing (NSUI for the INC and ABVP for the BJP) office bearers.

We built the database of Indian politicians using NivaDuck [39] [1] - an ML-based classification pipeline that we have developed to identify political actors on Twitter in a given country. It leverages Twitter profile description text and tweet content of known politicians to identify new politicians. So far, NivaDuck has identified over 18500 Indian politicians and over 8000 US politicians - the largest such archive to our knowledge. These accounts have been manually verified and annotated with their party and state. The most significant contribution of NivaDuck is its ability to find politicians that are not listed in official data sources like the Federal Election Commission (FEC) in the US or the Election Commission of India (ECI) in India, especially in the nations of the Global South.

We built this database by iteratively collecting the list of politicians, starting with a random sample of 1700+ manually curated Twitter handles of Indian politicians - members of parliament, state leaders and grass-root activists from 42 major national and state parties. We use these to

[1] Marathi word for 'selector'

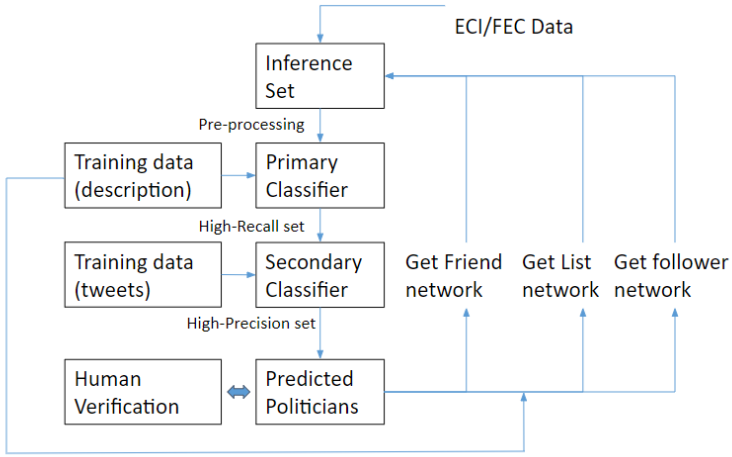

Fig. 1. NivaDuck's classification pipeline

train NivaDuck's classifiers. Figure 1 shows the two-stage classification pipeline. The primary classifier considers only the Twitter profile description text whereas the secondary classifier is trained on tweets. For both classifiers, we trained machine-learning Logistic Regression models [41] with unigrams, bigrams and trigrams of the profile description and tweet text[2] as feature vectors respectively. We used GridSearchCV to optimise for the regularization parameter and the precision-recall curve to select classification thresholds that yielded a high recall output from the primary classifier and a high precision set from the secondary classifier. Overall, NivaDuck had a precision score of 90 percent and recall score of 65 percent on the test set. We prioritized precision over recall to reduce false positives.

To find new politicians, we used four different sources - friend-network and list-network of known politicians, election commission database, and users who tweeted trending political hashtags. These accounts were fed to NivaDuck to identify new politicians. Every classified politician was manually verified to remove false positives. NivaDuck's precision on the predicted set of politicians varied between 85pc to 93pc, depending on the source. The limitations of this archive are threefold. First, it may exclude accounts that are not well networked to other politicians through friend/follower links. Second, we observed a bias toward politicians of the two national parties - BJP and INC - in finding new politicians. We mitigated this by manually adding missing politicians from major regional parties in our database. Lastly, given that it is trained to be precise, it may exclude politicians that have very few tweets and those whose tweets do not match the training sample.

We chose an ML-based procedure for three reasons. Firstly, per our knowledge, there are no large public repositories of social media handles of politicians for India. Prior work has only considered major parties, their senior politicians and other members of parliament. While the ECI publishes social media handles of candidates, these are often outdated as politicians change parties and do not account for those who have not contested national or state elections. Secondly, manual collection of these accounts is error-prone and tedious, making it hard to replicate over time. Moreover, the large, multilingual and multi-party Indian political system makes human effort even more inefficient and ineffective. An ML-based procedure allows for a large scale study like the one we pursued here. Thirdly, we intend to repeat this study for other large democracies, especially in the Global South.

[2]Tweet text was featurized using Google's Universal Sentence Encoder [7]

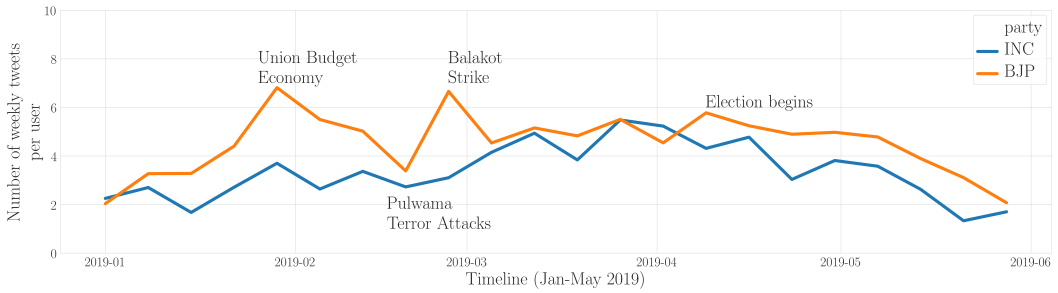

Fig. 2. Weekly tweet output per user by BJP and INC

NivaDuck's scalability and adaptability make it a suitable method to build large corpora of political figures on Twitter worldwide.

For all accounts thus collected, we manually added party and state annotations. We verified each handle and marked their state as follows: (1) if they were an elected representative, (2) if they were un-elected and had a location or reference to a location in their description or screen name (eg. @MumbaiCongress, @AAPKarnataka). Their location was marked accordingly. For the remaining handles we marked the location as 'Unknown'. For party labels, we used the Twitter profile description, screen name and the latest official ECI database of candidates in national and state elections.

For this study, we selected BJP and INC politicians from our database who had at least 100 followers. This yielded 4280 BJP politicians that posted 400724 original tweets and 3102 INC politicians that posted 208483 original tweets between Jan-May 2019 - the general election campaign period. We selected hashtags that appeared in at least 100 original tweets for our study. This produced a sample of 1208 unique hashtags, after merging those with different case structure (eg. #ModiAgain, #Modiagain, and #MODIAgain).

Figure 2 gives a broad view of weekly Tweet activity of BJP and INC politicians, normalized for the number of users. Notably, even on a daily basis, BJP posts more tweets than the INC ($T=5.449$, p-value < 0.001).

3.1 Hashtag typology

We identified four categories of issues for our analysis - `Nationalism`, `Development`, `Corruption` and `Religion`. While `Corruption` is a perennial issue of Indian politics [23], INC's #Chowkidar hashtags made it a focus of their campaign. `Religion` and `Nationalism` have been recurring campaign themes for the BJP since the 1990s [15, 52] and took center stage with controversies around the Ayodhya Ram Temple-Babri Mosque dispute and the Pulwama terror attacks and Balakot air-strikes respectively. As both parties aggressively promoted their social welfare proposals like the INC's NYAY and BJP's Ayushman Bharat, we added `Development` to our typology.

The first two authors independently encoded hashtags into one of the four categories, as per the definition in table 1. The categories were mutually exclusive. For hashtags that could be matched to multiple categories, the coders were asked to select the most suitable category. As an example, #ChowkidarNahiRozgarChahiye (We want jobs not gatekeeper) relates to both `Development` and `Corruption`[3], but was annotated as `Development` due to the direct mention of jobs. We labeled hashtags that did not fit into any category and those which could not be confidently classified / strongly associated into the typology as `Other`. The inter-coder reliability of our annotations,

[3]The phrase 'Chowkidar' (gatekeeper) has been used by both sides in relation to corruption allegations

Category	Definition	Examples
Nationalism	References to Pulwama, surgical strikes, Balakot, Pakistan, Indian armed forces, CRPF, martyrs or freedom fighters	#BJPFailedNationalSecurity, #BalakotAirStrike, #CRPFJawans, #CongressPakistanUnited
	Call to boycott 'anti-national' actors	#ExposeDeshDrohis (traitors)
Development	Govt. programme / policy proposal	#MakeInIndia, #NYAYforIndia
	Economy, health, unemployment, farmers	#Modinomics, #AyushmanBharat
Corruption	Reference to corruption related controversies	#ModiScamCentury, #RafaleGrandExpose
Religion	Issues, events about religion	#AyodhyaHearing, #INCMinorityConvention
	Religious festivals, personalities from religious / spiritual organisations	#HappyHoli, #ModiInKumbh, #JummaMubarak

Table 1. Definition of issue-based categories for hashtags used by BJP and INC in the 2019 election campaign

measured using Cohen's *kappa* statistic, was 0.72. The disagreements between the two coders were resolved by the third author. In all, we labelled 131 hashtags as `Nationalism`, 96 as `Development`, 55 as `Corruption` and 53 as `Religion`. There were 859 `'Other'` hashtags.

Apart from this typology, we also labeled each hashtag as 'BJP-related' or 'INC-related' if it contained a reference to the respective party, affiliated organisations or its politician(s).

4 ANALYSIS

4.1 Issue-based preferences of campaigns

We estimated the topical focus of the campaigns of BJP and INC using two methods. Firstly, we report the likelihood of the two parties using hashtags relating to our typology. We modeled the number of tweets of each category by each party on a daily basis as a binomial variable and used a mixed random effects model to estimate the odds that a party will post tweets about that category. The regression formula was as follows: [4]

$$cbind(N_{category}, N_{other}) \sim (1|day) + party + category + party * category \quad (1)$$

As we wanted to study the relative preference of the two parties in posting tweets about a particular subject, say `corruption`, we used log-odds ratio to model this metric. Table 2 shows that the BJP is more than twice as likely to tweet about `nationalism` and `religion` than the INC but less than half as likely to tweet with `corruption` related hashtags. On development, the difference between the parties is relatively small.

Secondly, we consider the total number of politicians that used these hashtags throughout the study period. Hemphill et al. [19] defined partisan score of each hashtag as the Chi-Squared statistic of dependence between number of users that used a hashtag and their party. Using their method, we report the mean partisan score for the four categories.

We categorized all hashtags within 0.5 standard deviation of the mean as 'non-partisan', between 0.5 to 1.5 standard deviations as 'leans-bjp/inc' and those beyond 1.5 standard deviations as 'strong-bjp/inc'. Figure 3 shows the mean partisan scores for the four types of hashtags. They are color-coded as per our typology and the respective bands indicate the mean partisan score of the four respective categories. The x-axis shows the partisan score on a log scale and partitions labeled at the top show

[4]We used R to model the regression [43]

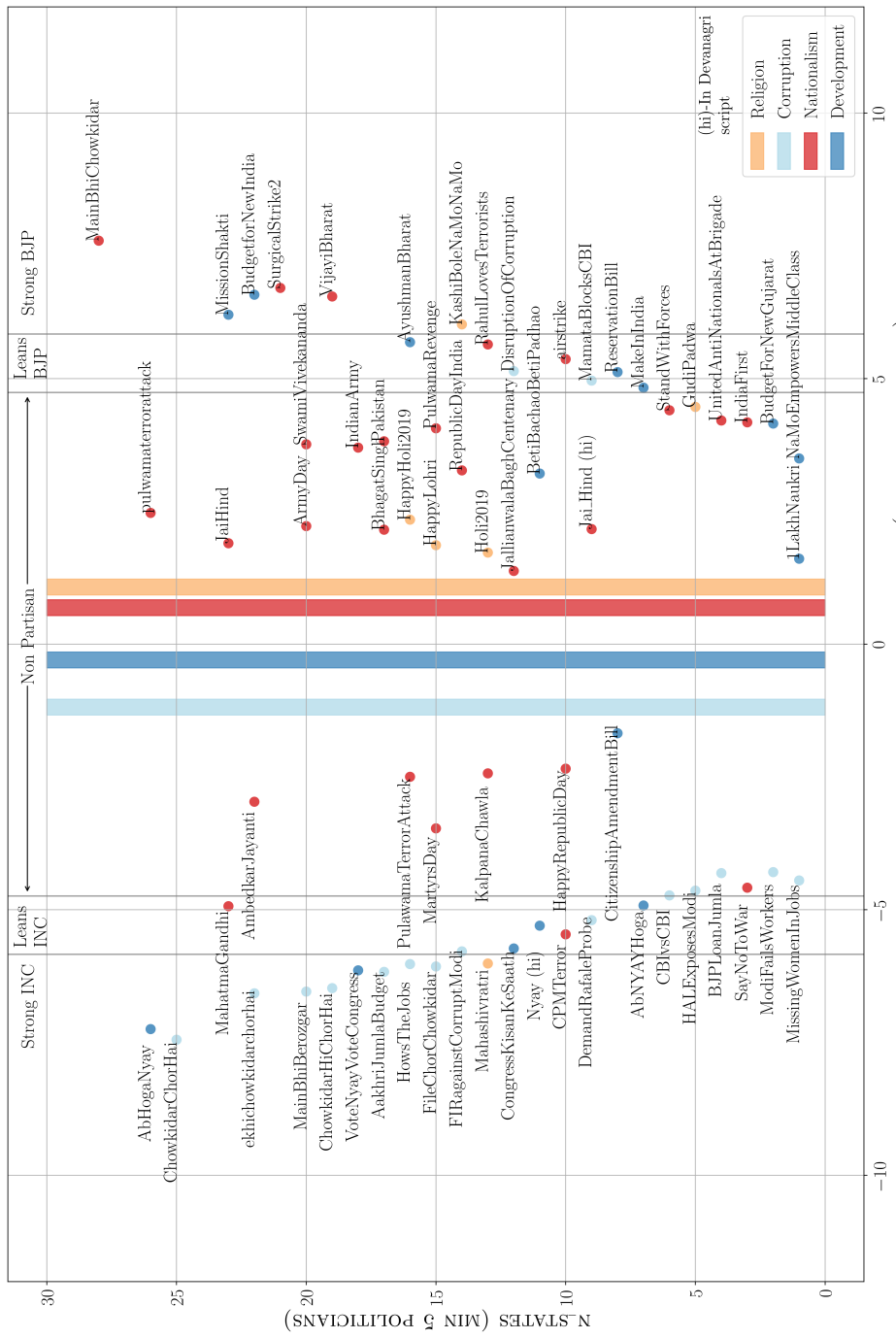

Fig. 3. Average partisanship scores for nationalism, religion, development and corruption (Bands indicate the mean log score of partisanship, y-axis plots the number of states that had at least 5 politicians using the hashtag)

Type	Odds ratio (BJP/INC)	Std. Err.	p-value	Total tweets posted BJP	INC
Corruption	0.419	0.00620	<0.0001	3964	12674
Development	1.114	0.00978	<0.0001	15787	12037
Nationalism	2.170	0.02044	<0.0001	10493	3907
Religion	2.511	0.05036	<0.0001	6923	2103
Self-Promotion	4.874	0.05648	<0.0001	237423	64223
Attacking-Opponent	0.484	0.05648	<0.0001	13824	52247

Table 2. Likelihood ratio of parties' typological preferences

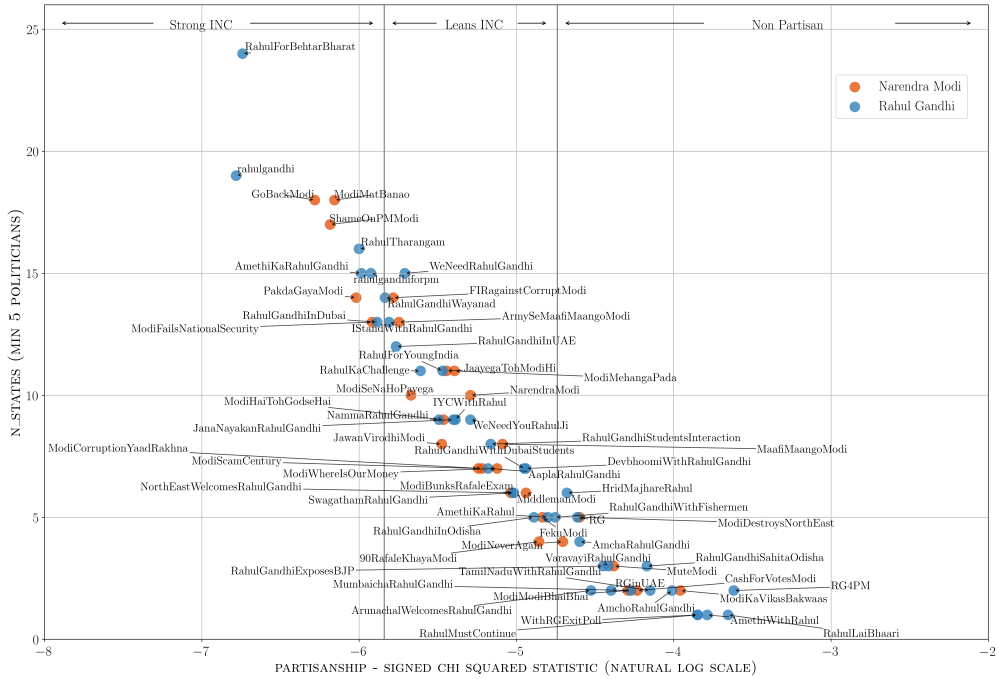

Fig. 4. Sample of hashtags with negative partisnaship scores that refer to Narendra Modi or Rahul Gandhi

the partisan category of the hashtag. The results corroborate those from table 2. For an interactive visualization of all hashtags, please see our GitHub repository listed in the appendix.

We then considered the references to 'BJP' and 'INC' by the two parties. We defined tweets with hashtags that referred to the same party as 'Self-promotion' and to the opposing party as 'Attacking-opponent'. The BJP is 4.8 times more likely to promote itself than the INC does while the odds that the INC will attack the BJP are more than twice the odds of BJP attacking INC. Within the parties, the BJP very rarely attacked the INC as against promoting itself (odds ratio = 0.11, p-value<0.0001) whereas the INC was more likely to attack than self-promote (odds ratio = 1.10, p-value<0.0001).

Figures 4 and 5 show samples of hashtags that referred to Modi and Rahul Gandhi used by the INC and BJP respectively. We included English hashtags that used direct references to their names

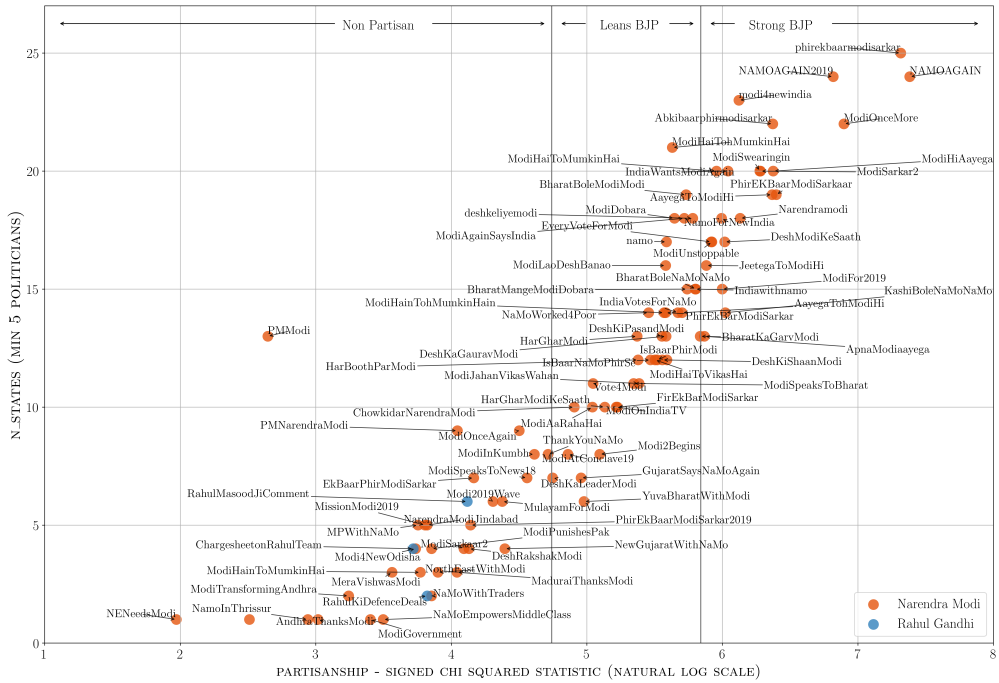

Fig. 5. Sample of hashtags with positive partisnaship scores that refer to Narendra Modi or Rahul Gandhi

or to fan-names like 'NaMo' and 'RG'. The BJP used twice as many leader-centric hashtags as the INC, with all of the 'Strong-BJP' hashtags being about Modi, while the 'Strong-INC' hashtags refer to both. This is reinforces our claim that the BJP promotes itself (and its leadership) way more than attacking its opponents, while the INC is split between the two.

4.2 Effectiveness of issue-specific messaging

We estimated the impact of differing categorical preferences of the two campaigns using two metrics - retweet count and trend value. The trend value was estimated as the cumulative number of hours that a given hashtag trended on that day. The sum of the hashtags' scores yielded the trend value of the tweet. We excluded non-partisan tweets from our analysis. We used a mixed effects model with the following regression formulae.

$$retweet\_count \sim (1|day) + (1|politician) + party + party * category + followers\_count + language + size \quad (2)$$

$$trend\_value \sim (1|day) + (1|politician) + party + party * category + followers\_count + language + size \quad (3)$$

Tables 3 and 4 list show that the INC achieves higher retweets than the BJP across all types of tweets. However, the BJP has a higher trend score for its tweets i.e. its hashtags trend more often than those of the INC. We excluded results that were not significant.

Table 3. Regression Results for retweets earned by BJP and INC politicians when they tweet about different types of hashtags

Dependent variable: log_retweet_count	
partyINC	0.261808^{***}
categoryDevelopment	-0.205201^{***}
categoryReligion	-0.117734^{***}
categoryOther	-0.081895^{***}
categoryNationalism	-0.150060^{***}
log_foll	0.462637^{***}
log_size	0.229264^{***}
partyINC:categoryReligion	0.198934^{***}
partyINC:categoryDevelopment	0.134827^{***}
Observations	229,556
Contrasts	Estimate
Corruption::BJP−INC	-0.262^{****}
Development::BJP−INC	-0.397^{****}
Nationalism::BJP−INC	-0.290^{****}
Other::BJP−INC	-0.309^{****}
Religion::BJP−INC	-0.461^{****}
Note:	*p<0.1; **p<0.05; ***p<0.01; ****p<0.0001

Table 4. Regression results for trend value of tweets by BJP and INC politicians when they tweet about different types of hashtags

Dependent variable: log_trend_value	
partyINC	-0.23281^{***}
categoryDevelopment	-0.217162^{***}
categoryOther	-0.24841^{***}
log_foll	-0.003289^{***}
log_size	0.023918^{***}
partyINC:categoryDevelopment	0.20975^{***}
partyINC:categoryNationalism	0.170008^{***}
partyINC:categoryOther	0.21528^{***}
Observations	229,556
Contrasts	Estimate
Corruption::BJP−INC	0.2328^{****}
Development::BJP−INC	0.0231^{****}
Nationalism::BJP−INC	0.0628^{****}
Other::BJP−INC	0.0175^{****}
Religion::BJP−INC	0.1752^{****}
Note:	*p<0.1; **p<0.05; ***p<0.01; ****p<0.0001

5 DISCUSSION

There are key differences between the two parties topical preferences that merit attention. The BJP's has many more tweets that focus on issues of `Religion` and `Nationalism`, which suggests that it does appeal to its traditional right-wing base, as opposed to the past, when the party, and its leadership made an explicit attempt to underplay religious or nationalistic tones in its campaign [37].

Also, we find that in BJP politicians' tweets, both 'Modi' and 'BJP' are much more widely used especially alongside terms that imply the nation such as 'Bharat' or 'India' - for instance – #modi4newindia, #VijayiBharat (victorious India), #BharatKaGarvModi (India's pride Modi) and #JitegaModiJitegaBharat (Modi wins, India wins).

In contrast, the INC framed `Corruption` through the viral #ChowkidarChorHai (Hindi - 'the gatekeeper is the thief'), in response to Modi's description of himself as the nation's 'Chowkidar' in 2014. Till Feb 2019, #ChowkidarChorHai trended online [24], but in what may be the BJP's most successful counter-campaign, it was able to turn the negative connotation of Chowkidar with "thief" into a term it owned. It did so by asking its followers to add the prefix "Chowkidar" to their names, and trend the #MainBhiChowkidar (Me too gatekeeper) hashtag [24, 42] starting mid-March 2019. The #MainBhiChowkidar campaign underlines the importance of a broad user-base since the retweet rates of specific messages may have some affective value, but the ability to move the discourse a notch in one direction is driven by strength in numbers.

The INC's strategy cedes the initiative to the BJP, potentially allowing it determine which topics get discussed and how they are framed. This is corroborated by the finding that the ruling party's messaging was almost entirely self-centered while the opposition was split between promoting itself and attacking the BJP. This has important implications for the nature of political discourse and who shapes it, not just in India, but for democracies across the world that have experienced personality centered right-wing parties gaining power on the backs of effective personality-driven social media outreach [17, 34, 54].

In conclusion, this research shows empirically what has often been argued anecdotally by commentators, that the election has seen a nationalistic shift by the BJP. It is important to situate this within the larger context of Indian politics in the last decade, since the BJP came to power eschewing its traditional right-wing Hindu-politics narrative and instead focusing on development and anti-corruption. As a party in power, attempting to defend itself, the turn to nationalism as well as religion underlines the party's recognition of emotional sectarian appeals as valuable at the votebanks. While there are nuances specific to Indian politics, the case of an incumbent, being attacked on corruption and development, turning to nationalistic and sectarian politics holds warnings for political movements worldwide. The successes in this lean towards a political enemy, rather than a policy offering, are portentous in an international political environment in which nation states increasingly see heightening amounts of polarization.

6 LIMITATIONS

In this work, we proposed a qualitative typology of issues focused by Indian politicians during the general elections period from January to May 2019. We further compared how effective were the politicians from the ruling and opposition parties in spreading the issue-specific messages we identified. The major limitation of our work is that it does not consider the text of the tweet. Our typology of issues and classification of tweets is based entirely on hashtags. The choice of hashtags stems from the critical networking affordance that hashtags provide [6, 48] and the practice of co-ordinated hashtag campaigns by political parties [24]. This method assumes convergence between the topical attributions of the tweet text and its hashtags. But some tweets with hashtags of a

specific category, say `Nationalism`, may not be related to the issue. For instance, classifying a tweet containing the hashtag "#CRPFJawans" (translates to Central Reserve Police Force Soldiers) as `'Nationalism'` could be wrong if someone used it in a different, non-political context. However, from the content of the tweets with this and other similar armed forces related hashtags, we know that such references by politicians are invariably a call to nationalistic sentiment. Second, 'meta-tweets' may include hashtags of the tweets they report, but cannot be labeled as tweets about that topic. Meta-tweets are tweets that talk about other tweets or just report statistics about trending hashtags or popular issues. Third, a tweet may contain strong language pertaining to a given subject without any hashtag used. Such tweets are not covered by our analysis.

7 FUTURE WORK

In the future, we would like to extend our work to address these limitations by using natural language inference techniques to analyse the content of tweets. One important extension could be to analyse the etymology and evolution of the different hashtags used by political parties in election campaigns. For instance, the INC's hashtags about Rahul Gandhi have greater liguistic diversity than the BJP's hashtags about Modi. Moreover, many hashtags are derived from rhymes, jingles and slogans used in Indian political campaigns for decades, and are now reverberating virtually through the affordances of social media. We also plan to study political speech in other large electoral systems by leveraging NivaDuck to compile a representative dataset of politicians for those regions. This would contribute to efforts in the CSCW community to understand collective communication campaigns on social media. In addition, this methodology can be extended to study partisan framing of key issues by politicians of specific locations, such as membership of the European Union in European countries, healthcare reform in the US, and the contentious Citizenship Amendment Act (CAA) in India.

ACKNOWLEDGMENTS

We would like to thank Ramgopal Chandresekaran, Mugdha Mohapatra, Faisal Lalani, Dibyendu Mishra, Azhagu Meena, and Zainab Akbar who contributed to compiling our database of Indian politicians. We are also grateful to Ashwin Rajadesingan for providing valuable inputs toward improving the content and writing of this paper.

REFERENCES

[1] Ana Adi, Kristofer Erickson, and Darren G Lilleker. 2014. Elite tweets: Analyzing the Twitter communication patterns of Labour party peers in the House of Lords. *Policy & Internet* 6, 1 (2014), 1–27.

[2] Eric PS Baumer, Mark Sueyoshi, and Bill Tomlinson. 2011. Bloggers and readers blogging together: Collaborative co-creation of political blogs. *Computer Supported Cooperative Work (CSCW)* 20, 1-2 (2011), 1–36.

[3] Leticia Bode, Alexander Hanna, Junghwan Yang, and Dhavan V Shah. 2015. Candidate networks, citizen clusters, and political expression: Strategic hashtag use in the 2010 midterms. *The ANNALS of the American Academy of Political and Social Science* 659, 1 (2015), 149–165.

[4] Javier Borge-Holthoefer, Walid Magdy, Kareem Darwish, and Ingmar Weber. 2015. Content and network dynamics behind Egyptian political polarization on Twitter. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. 700–711.

[5] Lia Bozarth and Joyojeet Pal. 2019. Twitter discourse as a lens into politicians' interest in technology and development. In *Proceedings of the Tenth International Conference on Information and Communication Technologies and Development*. ACM, 33.

[6] Axel Bruns and Jean E Burgess. 2011. The use of Twitter hashtags in the formation of ad hoc publics. In *Proceedings of the 6th European Consortium for Political Research (ECPR) General Conference 2011*.

[7] Daniel Cer, Yinfei Yang, Sheng-yi Kong, Nan Hua, Nicole Limtiaco, Rhomni St John, Noah Constant, Mario Guajardo-Cespedes, Steve Yuan, Chris Tar, et al. 2018. Universal sentence encoder. *arXiv preprint arXiv:1803.11175* (2018).

[8] Andrea Ceron and Giovanna d'Adda. 2016. E-campaigning on Twitter: The effectiveness of distributive promises and negative campaign in the 2013 Italian election. *New media & society* 18, 9 (2016), 1935–1955.

[9] Sunandan Chakraborty, Joyojeet Pal, Priyank Chandra, and Daniel M Romero. 2018. Political Tweets and Mainstream News Impact in India: A Mixed Methods Investigation into Political Outreach. In *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies*. ACM, 10.

[10] Hsia-Ching Chang. 2010. A new perspective on Twitter hashtag use: Diffusion of innovation theory. *Proceedings of the American Society for Information Science and Technology* 47, 1 (2010), 1–4.

[11] Pradeep K Chhibber and Susan L Ostermann. 2014. The BJP's fragile mandate: Modi and vote mobilizers in the 2014 general elections. *Studies in Indian Politics* 2, 2 (2014), 137–151.

[12] Jill P Dimond, Michaelanne Dye, Daphne LaRose, and Amy S Bruckman. 2013. Hollaback! The role of storytelling online in a social movement organization. In *Proceedings of the 2013 conference on Computer supported cooperative work*. 477–490.

[13] Yogesh K Dwivedi and KK Kapoor. 2015. Metamorphosis of Indian electoral campaigns: Modi's social media experiment. *International Journal of Indian Culture and Business Management* 11, 4 (2015), 496–516.

[14] Gunn Enli and Chris-Adrian Simonsen. 2018. 'Social media logic'meets professional norms: Twitter hashtags usage by journalists and politicians. *Information, Communication & Society* 21, 8 (2018), 1081–1096.

[15] Lars Tore Flåten. 2016. *Hindu nationalism, history and identity in India: Narrating a Hindu past under the BJP*. Routledge.

[16] Catherine Grevet, Loren G Terveen, and Eric Gilbert. 2014. Managing political differences in social media. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. 1400–1408.

[17] Max Hänska and Stefan Bauchowitz. 2017. Tweeting for Brexit: how social media influenced the referendum. (2017).

[18] Libby Hemphill, Aron Culotta, and Matthew Heston. 2013. Framing in social media: How the US Congress uses Twitter hashtags to frame political issues. *Available at SSRN 2317335* (2013).

[19] Libby Hemphill, Aron Culotta, and Matthew Heston. 2016. # Polar Scores: Measuring partisanship using social media content. *Journal of Information Technology & Politics* 13, 4 (2016), 365–377.

[20] Libby Hemphill, Jahna Otterbacher, and Matthew Shapiro. 2013. What's congress doing on twitter?. In *Proceedings of the 2013 conference on Computer supported cooperative work*. ACM, 877–886.

[21] Libby Hemphill and Matthew A Shapiro. 2019. Appealing to the base or to the moveable middle? Incumbents' partisan messaging before the 2016 US congressional elections. *Journal of Information Technology & Politics* (2019), 1–17.

[22] Jane Im, Eshwar Chandrasekharan, Jackson Sargent, Paige Lighthammer, Taylor Denby, Ankit Bhargava, Libby Hemphill, David Jurgens, and Eric Gilbert. 2019. Still out there: Modeling and Identifying Russian Troll Accounts on Twitter. *arXiv preprint arXiv:1901.11162* (2019).

[23] Rob Jenkins. 2007. India's Unlikely Democracy: Civil society versus corruption. *Journal of Democracy* 18, 2 (2007), 55–69.

[24] Faisal Lalani Joyojeet Pal, Anmol Panda. 2019. How #BJP fused with #StrongIndia in 2019. *Livemint* (2019). https://www.livemint.com/elections/lok-sabha-elections/how-bjp-fused-with-strongindia-in-2019-1557414405626.html

[25] Andreas Jungherr. 2016. Twitter use in election campaigns: A systematic literature review. *Journal of information technology & politics* 13, 1 (2016), 72–91.

[26] Juhi Kulshrestha, Motahhare Eslami, Johnnatan Messias, Muhammad Bilal Zafar, Saptarshi Ghosh, Krishna P Gummadi, and Karrie Karahalios. 2017. Quantifying search bias: Investigating sources of bias for political searches in social media. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. 417–432.

[27] Mirko Lai, Cristina Bosco, Viviana Patti, and Daniela Virone. 2015. Debate on political reforms in Twitter: A hashtag-driven analysis of political polarization. In *2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*. IEEE, 1–9.

[28] Ankit Lal. 2017. *India social: how social media is leading the charge and changing the country*. Hachette UK.

[29] Amparo López-Meri, Silvia Marcos-García, and Andreu Casero-Ripollés. 2017. What do politicians do on Twitter? Functions and communication strategies in the Spanish electoral campaign of 2016. *El profesional de la información* 26, 5 (2017), 795–804.

[30] David Gunnarsson Lorentzen. 2017. Is it all about politics? A hashtag analysis of the activities of the Swedish political Twitter elite. *Human IT: Journal for Information Technology Studies as a Human Science* 13, 3 (2017), 115–155.

[31] Ingunn Lunde. 2016. Hashtag Poetics: Political humour on Russian Twitter. *Zeitschrift für Slawistik* 61, 1 (2016), 102–118.

[32] Vidya Narayanan, Bence Kollanyi, Ruchi Hajela, Ankita Barthwal, Nahema Marchal, and Philip N Howard. 2019. News and information over Facebook and WhatsApp during the Indian election campaign.

[33] Taberez Ahmed Neyazi. 2019. Digital propaganda, political bots and polarized politics in India. *Asian Journal of Communication* (2019), 1–19.

[34] Brian L Ott. 2017. The age of Twitter: Donald J. Trump and the politics of debasement. *Critical studies in media communication* 34, 1 (2017), 59–68.

[35] Jahna Otterbacher, Libby Hemphill, and Matthew A Shapiro. 2012. Tweeting vertically? Elected officials' interactions with citizens on Twitter. In *CeDEM (Conference for E-Democracy and Open Government) Asia 2012*.

[36] Joyojeet Pal, Priyank Chandra, Padma Chirumamilla, Vaishnav Kameswaran, Andre Gonawela, Udit Thawani, and Pritika Dasgupta. 2017. Innuendo as Outreach:@ narendramodi and the Use of Political Irony on Twitter. *International Journal of Communication (19328036)* 11 (2017).
[37] Joyojeet Pal, Dinsha Mistree, and Tanya Madhani. 2018. A Friendly Neighborhood Hindu. In *CeDEM Asia 2018: Proceedings of the International Conference for E-Democracy and Open Government; Japan 2018*. Edition Donau-Universität Krems, 97.
[38] Joyojeet Pal, Udit Thawani, Elmer Van Der Vlugt, Wim Out, Priyank Chandra, et al. 2018. Speaking their mind: Populist style and antagonistic messaging in the tweets of Donald Trump, Narendra Modi, Nigel Farage, and Geert Wilders. *Computer Supported Cooperative Work (CSCW)* 27, 3-6 (2018), 293–326.
[39] Anmol Panda, A'ndre Gonawela, Sreangsu Acharyya, Dibyendu Mishra, Mugdha Mohapatra, Ramgopal Chandrasekaran, and Joyojeet Pal. 2020. NivaDuck - A Scalable Pipeline to Build a Database of Political Twitter Handles for India and the United States. In *Proceedings of the 11th International Conference on Social Media and Society (SMSociety '20)*. Association for Computing Machinery, New York, NY, USA, 15.
[40] Souneil Park, Minsam Ko, Jungwoo Kim, Ying Liu, and Junehwa Song. 2011. The politics of comments: predicting political orientation of news stories with commenters' sentiment patterns. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work*. 113–122.
[41] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12 (2011), 2825–2830.
[42] PTI. 2019. BJP steps up its 'Main bhi chowkidar' campaign. *Hindu Business Line* (2019). https://www.thehindubusinessline.com/news/bjp-steps-up-its-main-bhi-chowkidar-campaign/article26568670.ece
[43] R Core Team. 2019. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/
[44] Saqib Saeed, Markus Rohde, and Volker Wulf. 2011. Analyzing political activists' organization practices: findings from a long term case study of the european social forum. *Computer Supported Cooperative Work (CSCW)* 20, 4-5 (2011), 265–304.
[45] Bryan C Semaan, Scott P Robertson, Sara Douglas, and Misa Maruyama. 2014. Social media supporting political deliberation across multiple public spheres: towards depolarization. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. 1409–1421.
[46] Matthew A Shapiro and Libby Hemphill. 2017. Politicians and the policy agenda: Does use of Twitter by the US Congress direct New York Times content? *Policy & internet* 9, 1 (2017), 109–132.
[47] Betwa Sharma. 2016. Amit Shah To BJP Workers: 25,000 Followers On Social Media Or Else. *Huffington Post India* (2016).
[48] Sanjay Sharma. 2013. Black Twitter? Racial hashtags, networks and contagion. *New Formations* 78, 78 (2013), 46–64.
[49] Jieun Shin, Lian Jian, Kevin Driscoll, and François Bar. 2017. Political rumoring on Twitter during the 2012 US presidential election: Rumor diffusion and correction. *new media & society* 19, 8 (2017), 1214–1235.
[50] Tamara A Small. 2011. What the hashtag? A content analysis of Canadian politics on Twitter. *Information, communication & society* 14, 6 (2011), 872–895.
[51] Kumar Uttam. 2018. Get 3 lakh 'genuine' Facebook likes each, Modi tells BJP MPs. *Hindustan Times* (2018).
[52] Milan Vaishnav. 2019. The BJP in Power: Indian Democracy and Religious Nationalism.
[53] Yu Wang, Jiebo Luo, Richard Niemi, Yuncheng Li, and Tianran Hu. 2016. Catching fire via" likes": Inferring topic preferences of trump followers on twitter. In *Tenth International AAAI Conference on Web and Social Media*.
[54] Chris Wells, Dhavan V Shah, Jon C Pevehouse, JungHwan Yang, Ayellet Pelled, Frederick Boehm, Josephine Lukito, Shreenita Ghosh, and Jessica L Schmidt. 2016. How Trump drove coverage to the nomination: Hybrid media campaigning. *Political Communication* 33, 4 (2016), 669–676.

A ONLINE RESOURCES

We have included interactive visualizations of partisan scores all hashtags in our GitHub repository, along with other useful information about our methodology. Please find it here: https://github.com/anmolpanda/partisan_hashtags_india.

Received October 2019; revised January 2020; accepted March 2020