Twitter Discourse as a Lens into Politicians' Interest in Technology and Development

Lia Bozarth
University of Michigan
lbozarth@umich.edu

Joyojeet Pal
University of Michigan
joyojeet@gmail.com

ABSTRACT

We present a large-scale study of social media discourse in India focusing on politicians' tweet frequency of topics related to information technology, poverty, and development. Examining the Twitter feed of 477 political elites, we observe that politicians as a whole more frequently discuss development comparing to technology and poverty. In addition, national-party politicians holding higher offices are significantly more likely to discuss all three topics comparing to regional-party officials and non-elected politicians even when accounting for their Twitter account characteristics. Finally, politicians from different states demonstrate different topic priorities, suggesting differences in local political circumstances contribute to political elites' Twitter topic prioritization.

CCS CONCEPTS

• Human-centered computing → Social media; • Applied computing → Sociology; • Social and professional topics → User characteristics;

KEYWORDS

technology and development; social media; political communications; India politics; Indian policies

ACM Reference Format:
Lia Bozarth and Joyojeet Pal. 2019. Twitter Discourse as a Lens into Politicians' Interest in Technology and Development. In Proceedings of ICTD 2019 conference (ICTD). ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

In the past decade, political communication in much of the world has started shifting dramatically from journalist-mediated communications to direct political messaging as more politicians turned to social media such as Twitter and Facebook as their primary means of outreach. This trend has been much discussed in recent years [11, 21, 22]. Politicians in India have likewise begun to communicate directly on social media, carving their own brands and picking the issues they wish to discuss instead of mediating it through the mainstream press.

At the same time, the last few decades have also been critical to the shaping of a public discourse around technology as central to Indian vision of modern nationhood, where thought leaders have presented technology-driven development as an aspirational ideal [16, 27]. In addition to the visibility of an 'ICTD discourse' in the public - through its prominence in business circles, the academy, and the mainstream media, politicians have aggressively embraced these ideals and brought discussions of technology to their political messaging [28], both during electoral campaigns and more generally as part of their daily image-building exercises [24].

In this paper, we attempt to analyze the importance of technology and development as topics on politicians' Twitter feed in India. To do this, we seek keywords in politicians' messaging and categorize them as being about technology, development, and poverty (broadly categorized as ICTD related topics).

Our paper makes the following contributions:

- To the best of our knowledge, this is the first large-scale classification of tweets posted by a substantial number of Indian politicians. Using word2vec in combination with keywords matching, we are able to classify ICTD-related tweets with 85% precision and 82% recall.
- We show that while technology is a frequently appearing topic in their speech, on aggregate politicians more frequently tweet about development-related issues
- We show that politicians from national parties, which frequently need to present a pan-Indian vision, tweet about technology-related more than politicians from regional parties
- We show that politicians from different states tweet differently on technology and development, suggesting that certain states that have built an aspirational discourse around technology and that their politicians are likely to propagate with that vision online.

2 RELATED WORK

Despite early claims that social media could decentralize power and enabling more participative politics [10, 19], recent work has suggested that not only does social media enhance the control over public discourse that politicians and institutions can exercise [8, 23], but that it can even enhance political actors' ability to censor and monitor citizens [9, 12, 33].

ICTD, Jan 2019, Ahmedabad, India

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM.
https://doi.org/10.1145/nnnnnnn.nnnnnnn

In the West, prior studies have observed an increased effort from established politicians and candidates in utilizing social media to communicate directly with their constituents for the purpose of personal brand-building, campaigning, policy discussion, et cetera [5, 13, 17]. Studies focusing on the Global South observed a similar trend [20, 26] both in terms of politicians' individual strategies [3, 18] as well as the framing of issues [2, 7, 18]. Much work has noted the impacts on the democratic process as PR firms are increasingly central to managing the social media output of political organizations [29], particularly in the context of India [15, 25].

Many scholars [4, 14, 17] stress the importance of examining the impact of social media on political discourse and democracy at large. Luckily, the digital prints left by these political entities allow us to archive and topically examine tweets at unprecedented scale and depth [32]. Our paper aims to contribute to the studies of digital media and political communications by examining Indian politicians' Twitter discourse on two topics at scale. More specifically, we focus on a substantial number of Indian politicians with varied characteristics (e.g. party type, constituency, state) and assess how they tweet about topics related to technologies, development, and poverty (broadly categorized as ICTD-related topics) in a span of 4 years (2014-2018).

3 DATA

Our dataset consists of 1.49 million tweets in both English and Hindi [1] from Jan 2014 to August 2018, contributed by 484 distinct Twitter accounts of Indian politicians. Approximately 952K of the tweets are in English, while about 534K are in Hindi. To aggregate a comprehensive list of politicians, we compiled a list of all elected upper and lower houses of the Parliament, state chief ministers and cabinets, and major post-holders such as party presidents, general secretaries, searched for each name individually, and compiled a list of all their Twitter handles. We filter out all Twitter accounts that have fewer than 1000 followers or have posted fewer than 1000 tweets. The resulting 484 accounts span across 20 different parties with 414 accounts being in national parties (BJP, INC, CPI(M)) who supplied 1.35M tweets (or, 90.6% of total tweets) and 70 accounts are in regional parties with aggregated sum of 133K tweets (or, 9.4% of total tweets). Furthermore, 223 are in Lok Sabha, 69 in Rajya Sabha, 112 are state legislators, while the remaining 80 are cabinet members, party leaders, et cetera.

3.1 Tweet Classification

we first classify tweets into the following 4 categories: 1) information and communication technologies (TECHNOLOGY) related tweets, 2) poverty AND welfare (POVERTY) related tweets, 3) economics and development (DEVELOPMENT), and 4) baseline. The goal of selecting technology, poverty, and development is to consider the extent to which key topics of democratic relevance are part of the political communications of Indian politicians.

Keywords and/or Cosine similarity measurement based document clustering has been used extensively by many prior works [1, 6, 30] focused on political communication on social media. In this paper, we use a similar approach to cluster tweets into different issue-based and personal-appeal-based categories.

Classification procedure: We use word2vec and cosine similarity scoring to generate keywords for each category of non-baseline tweets, we then use these keywords to assign each tweet into the matching category. To be more specific, i) we first use gensim, an open source natural language processing (NLP) toolkit in python, to produce a 300 dimensional vector space representation of our corpus (here, each tweet is a single document). Within this space, each unique word is assigned a numeric vector. For instance, the word "govt" has the corresponding 300d vector [-1.9078784 1.654422 1.8524699...]. ii) Given that word vectors are positioned such that words that share similar contexts are located closer to each other in the vector space (i.e. the cosine of the angle between the vectors of a pair of semantically similar words are smaller), we are able to manually select a handful keywords related to TECHNOLOGY, such as "technology", "digital", et cetera, and then use cosine similarity scoring to discover additional keywords that share comparable semantic meanings to the selected words. iii) Using this approach, we generate 157 keywords for TECHNOLOGY, 91 for POVERTY, and 153 for DEVELOPMENT. iv) For each tweet, we categorize it as TECHNOLOGY if it contains 1 or more related keywords (similar for other categories). If a tweet has no matching keywords, it's assigned as a baseline. Additionally, a tweet can belong to multiple categories. The complete list of keywords can be found at https://lbozarth.github.io/ictd_keywords.pdf.

Using this approach, we label 51.3K or 3.4% tweets as TECHNOLOGY, 56.6K or 3.8% as POVERTY, and 107K or 7.2% as DEVELOPMENT. This suggests that politicians tweet more about development comparing to technology and poverty. In order to assess the performance of our classification, for each category of tweets, we select a representative sample (e.g. we calculate sample size using 95% confidence level and ± 3% confidence interval) and manually assess whether each tweet indeed focuses on technology, poverty, and/or development related topics. We see an 80.3% accuracy for TECHNOLOGY, 82.0% for POVERTY, and 90.0% for DEVELOPMENT. In addition, we aggregate all 3 categories of tweets together and observe a combined precision and recall of 85.4%, and 82.0% respectively.

4 ANALYSES

We first examine the distribution of contributions for each category of original tweets (i.e not including retweets) focusing on each party type. For TECHNOLOGY, POVERTY, and DEVELOPMENT tweets, politicians in national parties contributed a total of 47.3K (91.6%), 42.3K (94.0%), and

[1] We identify and filter out the small subset of tweets written in regional languages

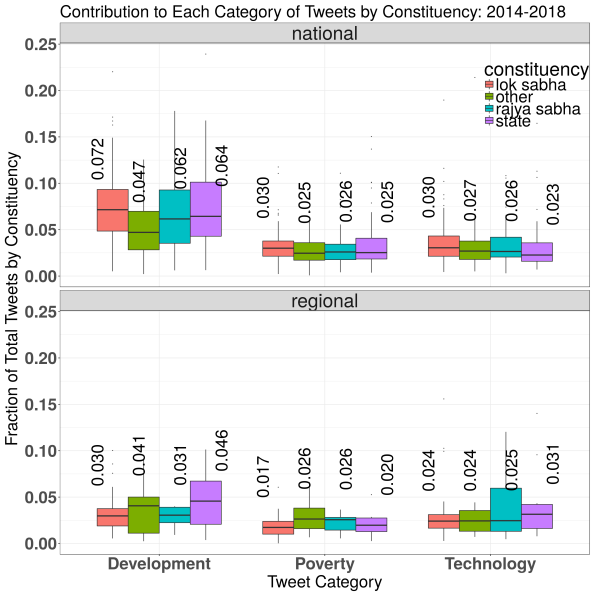


Figure 1: Distribution of Tweet Fraction by Tweet Category, Party Type, and Constituency. Note, the Labeled Numbers are the Median Tweet Fractions.

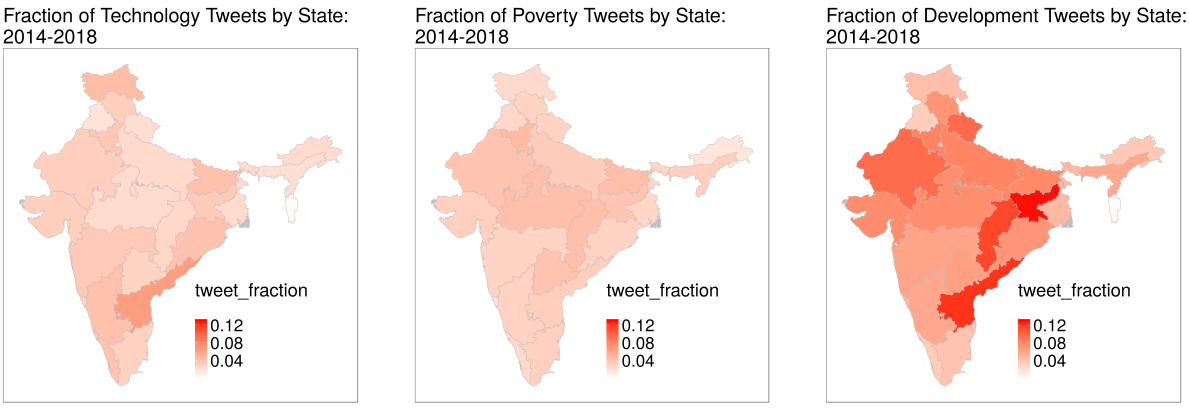
Table 1: Regression Result: Tweet Fraction by Tweet Category, Party Type, and Constituency

	Dependent Variable:
	tweet_fraction
POVERTY	−0.034***
	(0.002)
TECHNOLOGY	−0.030***
	(0.002)
Regional Party	−0.014***
	(0.002)
Constituency: other	−0.010***
	(0.002)
Constituency: rajya sabha	−0.003
	(0.002)
Constituency: state	−0.0003
	(0.002)
followers_count	0.001
	(0.001)
friends_count	−0.003*
	(0.001)
statuses_count	0.006***
	(0.002)
Constant	0.047***
	(0.005)
Observations	1,415
R^2	0.265
Adjusted R^2	0.260
Residual Std. Error	0.027 (df = 1405)
F Statistic	56.153*** (df = 9; 1405)
Note:	*p<0.1; **p<0.05; ***p<0.01

95.4K (97.4%) respectively while politicians in regional parties contributed 4.4K (8.4%), 2.7K (6.0%), 5.2K (2.6%). The median number of tweets contributed by politicians from national parties are 78, 77, and 169 for TECHNOLOGY, POVERTY, and DEVELOPMENT respectively. The numbers are 33, 26, and 48, for politicians from regional parties. These observations indicate that political elites are more focused on tweeting about development. Additionally, politicians in national parties are more actively tweeting about ICTD topics comparing to those from regional parties.

One possible explanation could be politicians in higher positions (e.g. being a Parliament member instead of a state legislator) are more inclined to post about ICTD-related tweets, and that regional-party politicians may be more inclined to tweet about other, possibly local, issues. Thus, we further breakdown politicians based on their constituency (i.e. Lok Sabha, Rajya Sabha, state-level legislature, or other) in addition to party type. As shown on Figure 1, we see that the Parliament legislators from national parties contributed a substantially higher fraction of tweets in DEVELOPMENT (median fraction of tweets is 7.2% for Lok Sabha, and 6.3% for Rajya Sabha) comparing to politicians who are also in the Parliament but come from regional parties (median fraction of tweets is 3.3% for Lok Sabha, and 2.6% for Rajya Sabha). Contributions to POVERTY and TECHNOLOGY by Parliament members from national and regional parties, however, are more comparable. Additionally, national-party state-level legislators tweet about development slightly more than their regional-party counterparts, though the difference is smaller (6.7% v.s. 5.4%). Finally, non-elected officials from both national and regional parties have comparable contributions for all 3 topics, indicating that their tweeting behaviour is less affected by party type.

So far, our overall findings suggest that politicians from both party types preferentially tweet about development comparing to technology and poverty. Additionally, politicians from national parties tweet more about these topics comparing to politicians from regional parties. To test for statistical significance, for each politician i, we write the dependent variable as the fraction of total tweets by i for each topic (e.g. we count the number of DEVELOPMENT tweets by i and divide it by i's total number of tweets in our dataset). We also derive the independent variables: i) tweet category (TECHNOLOGY, POVERTY, and DEVELOPMENT), ii) party type (national, regional), iii) constituency, and finally we control for i's number of (iv) friends, (v) followers, and (vi) total tweets. Regression results are shown in Table 1. Comparing to development, politicians tweet approximately 3% less about technology and poverty. Additionally, regional-party and non-elected politicians tweet less about ICTD-related topics by roughly 1%. These observations are consistent with our prior conclusion. More interestingly, we also see that i's Twitter account characteristics explain very little variance in the model. In fact, we see that the correlation between tweet fraction and number of followers is statistically insignificant (e.g. having more followers is not correlated with tweeting more about ICTD-related topics).



(a) Technology (b) Poverty (c) Development

Figure 2: Fraction of Tweets For each Category by State

4.1 State-level Contribution Distribution

Thus far, we have examined the distribution of contributions to ICTD-related tweets by politicians focusing on the politicians' party type and constituency. Yet, India is one of the largest countries in the world with many states which have varied issues and issue priorities. For instance, Bihar, Chhattisgarh, Jharkhand are some of the poorest states in India line [31], whereas states such as Karnataka and Telangana are relatively wealthier and have major urban technology industry hubs of Bangalore and Hyderabad respectively.

For each state, we calculate the fraction of total tweets contributed by politicians from that state in each Tweet category. As shown in Figure 2, unsurprisingly more than 6% of tweets from Andhra Pradesh talked about technology whereas the Hindi belt region - comprising Rajasthan, Haryana, Madhya Pradesh, Bihar, and Chhattisgarh focused more on poverty. This can be attributed to the significant attention to issues of technology and development in particular by politicians in Andhra Pradesh, which has been known for its pro-technology governments under its Chief Minister N Chandrababu Naidu from the Telugu Desam Party (TDP). There has been much past work on Naidu and the centrality of both technology and development in his outreach efforts, our data confirms this as significant in the Twitter output.

We also find that the development are important in Chhattisgarh and Jharkhand, two traditionally poor states that have had governments that were fairly aggressive about discussing development in their public speeches.

Finally, on the aggregate level, we again see that politicians as a whole are far more concerned with development comparing to technology and poverty.

5 DISCUSSION

In this paper, we first applied large scale classification of tweets posted by a substantial number of Indian politicians. We then assessed these political elites tweet contribution to ICTD-related topics focusing on 3 key characteristics: i) party type, ii) constituency, and iii) state.

We demonstrated that politicians as a whole tweeted more about development comparing to technology and poverty (with parliament members from national parties being the most active). Indeed, some of these issues are intuitive – development is broader and talked about outside of poverty alone, whereas the core issue of poverty tends to be more important in states where it is a critical electoral issue. We also showed that politicians from different states placed emphasis on different topics. In other words. We illustrated that Indian politicians' social media strategies are not uniform.

There are several caveats in our study worth noting here. First of all, while we provided assessment of politicians' Twitter priorities our paper did not cover to what extent politicians are successful when utilizing these different strategies. Future work should focus on building normative measurements of politicians' success on Twitter. One possible metric could be how many new followers, retweets or mentions politicians gain for different type of issues. Second, our dataset spans a time period which includes election cycles for some states but not others, but we did not distinguish politicians' social media strategies from these different time periods (e.g. before, during, and after elections). Future work that makes this distinction can provide valuable insights into how politicians and politically inclined individuals shift their personal branding within different political environments. Finally, we emphasis that our goal here is not to suggest that Twitter data can be an indicator of what governments are likely to do, what we show here is that social media can offer a useful lens into how topic analyses can illustrate what politicians suggest the priorities for states may be.

REFERENCES

[1] L.A. Adamic and N. Glance. 2005-08. The political blogosphere and the 2004 US election: divided they blog. In Proceedings of the 3rd international workshop on Link discovery. 36–43.

[2] S. Ahmed, , and K.Jaidka J. Cho. 2016. The 2014 Indian elections on Twitter: A comparison of campaign strategies of political parties. Telematics and Informatics 33, 4 (2016), 1071–1087.
[3] S.and M.M.Skoric Ahmed. 2014. My name is Khan: the use of Twitter in the campaign for 2013 Pakistan General Election. In 47th Hawaii International Conference on System Sciences (HICSS), IEEE.
[4] W.L. Bennett and B. Pfetsch. 2018. Rethinking political communication in a time of disrupted public spheres. Journal of Communication 68, 2 (2018), 243–253.
[5] B. Bimber. 2014. Digital media in the Obama campaigns of 2008 and 2012: Adaptation to the personalized political communication environment. Journal of Information Technology and Politics 11, 2 (2014), 130–150.
[6] M. Conover, J. Ratkiewicz, M.R. Francisco, B. Gonçalves, F. Menczer, and A. Flammini. 2011. Political polarization on twitter. Icwsm 133 (2011), 89–96.
[7] J.C.and J.E.Camargo Correa. 2017. "Ideological consumerism in colombian elections. Cyberpsychology, Behavior, and Social Networking (2017).
[8] N. Curato. 2017. Flirting with authoritarian fantasies? Rodrigo Duterte and the new terms of Philippine populism. Journal of Contemporary Asia 47, 1 (2017), 142–153.
[9] Y.and C.Baden David. 2018. Reframing community boundaries: the erosive power of new media spaces in authoritarian societies. Information, Communication and Society (2018), 1–18.
[10] W.H. Dutton. 2009. The fifth estate emerging through the network of networks. Prometheus 27, 1 (2009), 1–15.
[11] G.S.and E.Skogerbø Enli. 2013. Personalized campaigns in party-centred politics: Twitter and Facebook as arenas for political communication. Information, Communication and Society 16, 5 (2013), 757–774.
[12] D. Gayo-Avello. 2017. Social Media Won't Free Us. IEEE Internet Computing 21, 4 (2017), 98–101.
[13] 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.
[14] P.N. Howard. 2006. New media campaigns and the managed citizen. Cambridge University Press.
[15] C. Jaffrelot. 2015. Narendra Modi and the power of television in Gujarat. Television and New Media (2015).
[16] D.and R.Ramamurti Kapur. 2001. India's emerging competitive advantage in services. Academy of Management Perspectives (2001), 15–2.
[17] D. Kreiss. 2016. Seizing the moment: The presidential campaigns' use of Twitter during the 2012 electoral cycle. New Media and Society 18, 8 (2016), 1473–1490.
[18] C. Lee. 2017. Facebooking to Power: The Social Media Presence of Malaysian Politicians. ISEAS Perspectives 74 (2017).
[19] B.D.and D.Mercea Loader. 2011. Networking democracy? Social media innovations and participatory politics. Information, Communication and Society 14, 6 (2011), 757–769.
[20] M Lufkens. 2016. Twiplomacy Study 2016. Burson-Marsteller, Editor, Geneva (2016).
[21] S.C. McGregor. 2018. Personalization, social media, and voting: Effects of candidate self-personalization on vote intention. new media and society 20, 3 (2018), 1139–1160.
[22] L. Meeks. 2016. Gendered styles, gendered differences: Candidates' use of personalization and interactivity on Twitter. Journal of Information Technology and Politics 13, 4 (2016), 295–310.
[23] B. Ohm. 2015. Organizing Popular Discourse with and against the Media: Notes on the Making of Narendra Modi and Recep Tayyip Erdogan as Leaders-without-Alternative. Television and New Media 16, 4 (2015), 370–377.
[24] J. Pal. 2017. The Technological Self in India: From Tech-savvy Farmers to a Selfie-tweeting Prime Minister. In Proceedings of the Ninth International Conference on Information and Communication Technologies and Development.
[25] J. Pal, P. Chandra, and V.V. Vydiswaran'. 2016. Twitter and the rebranding of Narendra Modi. Economic and Political Weekly 51, 8 (2016), 52–60.
[26] J.and A.Gonawela Pal. 2016. Political social media in the global South. Conference on e-Business, e-Services and e-Society. Springer.
[27] C.K.and A.Hammond Prahalad. 2002. Serving the world's poor, profitably. Harvard business review (2002), 80–9.
[28] G.K. Reddy. 2002. New populism and liberalisation: regime shift under Chandrababu Naidu in AP. Economic and Political Weekly (2002), 871–883.
[29] Tworzecki H Semetko, H. 2017. The fourth era of political communication. The Routledge Handbook of Elections, Voting Behavior and Public Opinion (2017).
[30] Stefan Stieglitz and Linh Dang-Xuan. 2013. Social media and political communication: a social media analytics framework. Social Network Analysis and Mining 3, 4 (2013), 1277–1291.
[31] Wikipedia contributors. 2018. List of Indian states and union territories by poverty rate — Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=List_of_Indian_states_and_union_territories_by_poverty_rate&oldid=850092303 [Online; accessed 20-September-2018].
[32] J. Wilkerson and A. Casas. 2017. Large-scale computerized text analysis in political science: Opportunities and challenges. Annual Review of Political Science 20 (2017), 529–544.
[33] J.L. Zittrain, R. Faris, H. Noman, J. Clark, , and C.Tilton R. Morrison-Westphal. 2017. The Shifting Landscape of Global Internet Censorship.