Election Trolling: Analyzing Sentiment in Tweets during Pakistan Elections 2013

Arjumand Younus[*†◇], M. Atif Qureshi[*†◇], Muhammad Saeed[†], Nasir Touheed[†], Colm O'Riordan[*], Gabriella Pasi[◇]
[*]Computational Intelligence Research Group, Information Technology, National University of Ireland, Galway, Ireland
[◇]Information Retrieval Lab, Informatics, Systems and Communication, University of Milan Bicocca, Milan, Italy
[†]Web Science Research Group, Faculty of Computer Science, Institute of Business Administration, Karachi, Pakistan
ayounus, maqureshi@iba.edu.pk, saeed@uok.edu.pk, ntouheed@iba.edu.pk, colm.oriordan@nuigalway.ie, pasi@disco.unimib.it

ABSTRACT

The use of Twitter as a discussion platform for political issues has led researchers to study its role in predicting election outcomes. This work studies a much neglected aspect of politics on Twitter namely "election trolling" whereby supporters of different political parties attack each other during election campaigns. We also propose a novel strategy to detect terms that are usually not associated with sentiment but are introduced by supporters of political parties to attack the opposing party. We demonstrate a lack of political maturity as evidenced through high percentage of political attacks in a developing region such as Pakistan.

Categories and Subject Descriptors

H.3.5 [**Information Storage and Retrieval**]: Online Information Services; J.4 [**Social and Behavioral Sciences**]: Sociology

Keywords

trolling, sentiment, political attacks

1. INTRODUCTION

Twitter has assumed the role of a political platform where a majority of users engage in political activism. Furthermore, many Twitterers express their political affiliations during elections and a number of studies have attempted to predict election results based on Twitter messages [2]. This paper studies a much neglected aspect of political activism during elections namely *"election trolling"* and our focus is to study elections in Pakistan which qualifies as a region where democracy has never reached a stable state[1]. During the recently held "Pakistan Elections 2013" a new kind of activism emerged namely *"trolling"* which was used to refer to the act of attacks by supporters of one political party against supporters of the opposing political party[2].

In an attempt to analyse the trolling phenomenon during "Pakistan Elections 2013", we perform sentiment analysis of the Twitter conversations between potential voters of various political parties. We also build a custom lexicon of "political slang" introduced by supporters of various political parties with the main intent of attacking opponents. Our approach utilizes a seed set of negative sentiment words and uses a collocation-based methodology to further detect words that are traditionally not associated with political sentiment but classify as political insults.

2. METHODOLOGY

Political Conversations on Twitter: We gather Twitter conversations between potential voters of various political parties through the Twitter API. We first extract tweets containing the names of various political parties[3]; this is then followed by fetching all the public tweets of Twitterers that engage in a significant amount of political discussions. From within the extracted public tweets of political activists, we then extract threads/conversations related to elections and the Twitterers engaged in these conversations are manually annotated with their political affiliations (i.e., the political party supported by them)[4]. We considered a total of 11,024 tweets comprising 3„451 conversations between 692 unique Twitterers. The average follower count of the Twitterers in our dataset is 56. The annotations enable us to determine the level of "trolling" exhibited by the supporters of various political parties. The set of Twitterers

[1]A large percentage of the Pakistani general population supports military dictatorship and considers democracy to be a worse form of government when compared with dictatorship.
[2]Most of the times these attacks are of a personal nature and qualified as ad hominem `https://en.wikipedia.org/wiki/Ad_hominem`
[3]Pakistan is a region with a multi-party system.
[4]The annotation is performed by a political expert who regularly keeps track of political happenings in the region.

WWW'14 Companion, April 7–11, 2014, Seoul, Korea.
ACM 978-1-4503-2745-9/14/04.
http://dx.doi.org/10.1145/2567948.2577352.

Twitterer	Tweet	Scores
A	@B Yeh...#Shame with fake account, this is how PTIans think they will get votes	shame(-3),fake(-2)
B	@A Stop making a fuss and fuck off.	fuss(-2),fuck(-3)
A	@B A dumb leader like IK can produce followers like you.	dumb(-2)
B	@A A corrupt leader like Noora can hire paid trolls like you.	corrupt(-3),Noora(-5)

Table 1: Example of Twitter Conversation between Two Political Opponents *A* and *B*

Political Party	Measures		
	tf	idf	$tf-idf$
PTI	0.55	0.45	0.75
PMLN	0.40	0.65	0.85
PPP	0.25	0.20	0.75
MQM	0.35	0.55	0.85
JI	0.15	0.15	0.65

Table 2: P@20 Measure for our Proposed Metrics across Various Political Parties

engaged in political conversations during "Elections 2013" is denoted by U_p.

Mining Sentiment in Political Conversations: The sentiment analysis methodology employed in this paper utilizes the SentiStrength lexicon [1] in addition to a custom-built "political slang" lexicon. We first measure troll activity of each Twitterer involved in political conversations as follows

$$Troll_{Activity} = \frac{f_{u,w} * |w_{score}|}{\sum_{u \varepsilon U_p} f_{u,w}}$$

where $f_{u,w}$ is the frequency of usage of a negative word w by a particular Twitterer, and $|w_{score}|$ is the sentiment score associated with a word. Note that the list of negative words along with corresponding scores are taken from the SentiStrength lexicon. We then use the vocabulary of Twitterers whose measure of troll activity exceeds some threshold; from this vocabulary we calculate the following values for the terms: *idf*, *tf-idf*, and *tf*. We then select 20 terms that have highest values of *idf*, *tf-idf*, and *tf* and these comprise our "political slang" lexicon. The calculations are performed separately for each political party so as to identify separately the terms each party uses to attack the other parties.

3. EVALUATION AND RESULTS

This section describes our evaluation methodology for the custom lexicon of political insult terms along with presenting an analysis of trolling activity during political conversations on Twitter.

Evaluation of Political Attack Terms: We ask three annotators to label each identified political attack term from the top 20 discovered terms of the previous step. This labeling was performed corresponding to each political party, and comprises assignment of a label to mark as relevant a term that denotes a political attack. Table 1 shows results of the evaluation and the values shown correspond to P@20 measure for the identified top 20 terms. It is clear from the evaluation that the *tf-idf* metric performs best when extracting political attack terms from a given tweet corpus.

Analysis of Political Conversations over Twitter: We now present an analysis of the amount of negative sentiment and political attacks with Twitter conversations of various political parties. We devise a scoring scheme to quantitatively measure the amount of trolling activity by supporters of various political parties. For each conversation pair, we use the SentiStrength lexicon for term matching of positive and negative words while using the scores within the lexicon to compute the final score of a conversation. Table 2 illustrates an example of a conversation with scores shown; note that the term "Noora" in last row has been detected as a political attack term used by PTI supporters and we assign a score of -5 to all the terms within our lexicon. The entire conversation of Table 2 is assigned a score of -20.

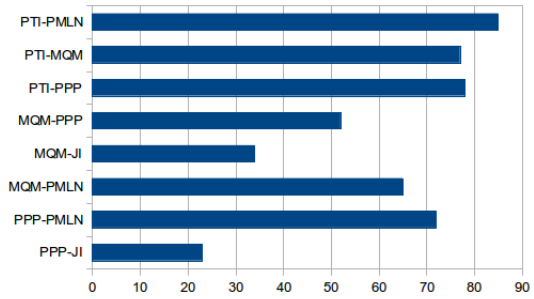

Figure 1: Percentage of Conversations with Negative Scores (Sentiment)

Figure 1 shows the percentage of conversations with negative scores for the various pairs of political parties involved in our analyses. The negative scores imply usage of negative sentiment words in addition to political attacks within the Twitter conversations and surprisingly this phenomenon is observed at a very high level for almost all pairs of political parties. The most aggressive behavior is demonstrated by supporters of PTI (Pakistan Tehreek-e-Insaaf) and PMLN (Pakistan Muslim League-Nawaz); and this is supported by news reports of constant tension between the two parties during election campaigns in 2013.

4. CONCLUSION

Our analysis confirms the hypothesis of a lack of political maturity in developing regions such as Pakistan. As future work, we aim to determine the extent to which supporters of various political parties focus on issues of national interest.

5. REFERENCES

[1] M. Thelwall, K. Buckley, G. Paltoglou, D. Cai, and A. Kappas. Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, 61(12):2544–2558, 2010.

[2] A. Tumasjan, T. O. Sprenger, P. G. Sandner, and I. M. Welpe. Predicting elections with twitter: What 140 characters reveal about political sentiment. *ICWSM*, 10:178–185, 2010.