Helping Citizens to Take Action on Political and Social Causes

Sindhu Vempati
West Virginia University
sv0014@mix.wvu.edu

Sindhura Athuluri
West Virginia University
sa0060@mix.wvu.edu

Srinivas Garapati
West Virginia University
sg0032@mix.wvu.edu

ABSTRACT
People tweet a lot on twitter, Twitter users are estimated to generate nearly 100,000 tweets every 60 seconds. All these tweets leads to Big data. Till date analysis says that, there is no particular use of those tweets. So, we got motivated to organize this big data to help the political and social causes. The platform will allow people to indicate which issues they find most important in order to receive information about what purchases to make and actions to take to exert their influence. The platform will use social media data, as well as public information from congress to make suggestions to citizens. People don't have complete idea on how to take action for a political cause. It is difficult to mobilize people to take action for a political cause (vote PLZ had less than half of the turnout it expected.). So by creating small micro tasks and suggesting citizens to take actions on it. Our Goal is to mobilize people to take micro actions for political causes. We will probe especially by mobilizing citizens to pay for items that benefit a particular politician. For example, buying a book from Trump university.

Author Keywords
Crowd sourcing; Twitter Bot; Tweets; Big Data; Hashtags; Crawlers;

1.INTRODUCTION

The election candidates in the previous trends used to employ people for the surveys of the election results to know their status of winning, but after the social media has emerged the social media trends are being considered[1] as the survey preliminary results. This social media trends are mainly helpful since it does not involve direct contact with people and could be done remotely. For this a software application that runs automated tasks called bots are being employed. Bots can also be said as social media accounts that automate interaction with other users. Bots can rapidly deploy messages to the users, duplicate themselves and can interact with users. There are various types of bots that could be deployed and in any area but they are mainly used in political campaigns, social issues, network bots, social media bots, data collections bots etc.

Networks of such bots are called botnets, a term combining robot with networks and describing a col- lection of connected computers with programs that commu- nicate across multiple devices to perform some task. There are legitimate botnets, like the Carna botnet, which gave us our first real census of device networks, and there are mali- cious botnets, like those that are created to launch spam and distributed denial-of-service (DDoS) attacks and to engineer theft of confidential information, click fraud, cyber-sabotage, and cyberwarfare.

Political bots are mainly employed to get the election results survey using the social platforms. Social bots are employed to cause an effect in the society using social media. Similarly, all these bots use social media data to collect the data and analyze or interact with the users. Generally, bots are dealt with large data set elements. The common data set elements mainly are social media data, political data, company's data, and other data sets. Bots need to clean and model diverse types of large datasets and handling high velocity of social media data. For example, Twitter users are estimated to generate nearly 100,000 tweets every 60 seconds and this data must be handled by the bots for the correct us- age.

Bots are very flexible, easy to deploy and are very cheap. Bots could be used in any area that has a scope for interacting with users or data collections. Often bot profiles lack basic account information such as screen names or profile pictures. Such accounts have become known as Twitter eggs because the default profile picture on the social media site is of an egg. While social media users get access from front-end web- sites, bots get access to such websites directly through a code to-code connection, mainly through the sites wide open ap- plication programming interface (API) that enables real-time posting and parsing of information. Bots could also be used in a negative way where bots could be deployed for internet attacks like sending spams, viruses and DDOS attacks can also be passed through bots. Politicians and political causes are generally backup by companies. Through purchases of items belonging to these companies, citizens can take direct action to support or oppose a political candidate. People don't have complete idea on how to act for a political cause. It is diffi- cult to mobilize people to act for a political cause. We can crowdsource citizens to take micro-actions whether to support or oppose a political candidate or cause. Crowdsourcing is the process of getting work or funding, usually online, from a crowd of people. The word is a combination of the words 'crowd' and 'outsourcing'. The idea is to take work and outsource it to a crowd of workers.

The social medium used in this platform is Twitter. Twitter provides free access to a sample of the public tweets posted on the platform. Twitter's precise sampling method is not known, but according to Twitter, the data available through the Streaming API is at most one percent of the overall global public communication on the platform at any given time. In order to get the most complete and relevant data set, the tweets were collected by following particular hashtags that

1.1 Twitter

Twitter is an online news and social networking service where users post and read short 140-character messages called ¨tweets¨. Registered users can post and read tweets, but those who are unregistered can only read them. Users access Twitter through the website interface, SMS or mobile de- vice app.In 2013, it was one of the ten most-visited websites and has been described as "the SMS of the Internet". As of March 2016, Twitter had more than 310 million monthly ac- tive users. On the day of the 2016 U.S. presidential election, Twitter proved to be the largest source of breaking news, with 40 million tweets (Bigdata) sent by 10 p.m. that day.

2. RELATED WORKS

The bots that are currently employed are mostly related to po- litical suggestions. Political bots have been particularly active on public policy issues, political crises, and elections.

2.1 Twitter Bots Shaping the Election

The related bots were employed only to collect the percent- age of supporters[2] for each candidate whereas the platform we are developing uses micro actions from the users. All the social media numbers, polls, or statistics are only as viable as the tools used to get to them. Political campaigns world- wide now use bots, software developed to automatically do tasks online, as a means for gaming online polls and artifi- cially inflating social-media traffic. [3]Recent analysis by the research team at Oxford University reveals that more than a third of pro-Trump tweets and nearly a fifth of pro-Clinton tweets between the first and second debates came from auto- mated accounts, which produced more than 1 million tweets in total. This data corroborates recent reports suggesting that both candidates' social media followings are highly au- tomated. [4]The team monitors political-bot activity around the world. They had data on politicians, government agencies, hacking collectives, and militaries using bots to disseminate lies, attack people, and cloud conversation. The widespread use of political bots solidified polarization among citizens.

2.2 Crowdsourcing for public policy and government

Though there is nothing qualitatively new about involving more people in government and policy processes, digital technologies in principle make it possible to increase the quantity of such involvement dramatically, by lowering the costs of participation and making it possible to tap into people's free time. [5]This difference in quantity is arguably great enough to obtain a quality of its own. We can thus be justified in using the term crowdsourcing for public policy and government to refer to new digitally enabled ways of involving people in any aspect of democratic politics and government, not re- placing but rather augmenting more traditional participation routes such as elections and referendums. [6]The growing interest in crowdsourcing for government and public policy must be understood in the context of the contemporary malaise of politics, which is being felt across the democratic world. Besides [7]allowing citizens to verify the outcome of a process, crowdsourcing can also be used to lend an air of inclusiveness and transparency to a process itself. This process legitimacy can then indirectly [8]legitimate the outcome of the process as well.

3. IMPLEMENTATION

3.1 Twitter API

In computer programming, an Application Programming In- terface (API) requires a software component in terms of its operations, their inputs and outputs and underlying types. Its main purpose is to define a set of functionalities that are in- dependent of their respective implementation, allowing both definition and implementation to vary without compromising each other. An application-programming interface (API) is a set of programming instructions and standards for access- ing a Web-based software application or Web tool. A soft- ware company releases its API to the public so that other software developers can design products that are powered by its service. In addition to accessing databases or com- puter hardware, such as hard disk drives or video cards, an API can be used to ease the work of programming graphical user interface components, to allow integration of new fea- tures into existing applications (a so-called" plug-in API"), or to share data between otherwise distinct applications. In practice, many times an API comes in the form of a library that includes specifications for routines, data structures, ob- ject classes, and variables. In some other cases, notably for SOAP and REST services, an API comes as just a specifi- cation of remote calls exposed to the API consumers. Twit- ter bases its Application Programming interface (API) of the Representational State Transfer (REST) Architecture. REST architecture refers to a collection of network design princi- ples that define resources and ways to address and access data. The architecture is a design philosophy, not a set of blueprints there's no single prearranged arrangement of com- puters, servers and cables. For Twitter, a REST architecture in part means that the service works with most Web syndi- cation formats. Using twitters API we can answer questions like: How many friends/followers do I have? Who am I fol- lowing that is not following me back? Who is following me that I am not following back? Who are the friendliest and least friendly people in my network? Who are my mutual friends? Given all of my followers and all of their followers, what is my potential influence if I get retweeted?

3.2 RESTFul and OAuth:

REST stands for Representational State Transfer. It is a client-server, communications protocol and in virtually all cases, the HTTP protocol is used. REST is an architecture style for designing networked applications. The idea is that, rather than using complex mechanisms such as CORBA, RPC or SOAP to connect between machines, simple HTTP is used to make calls between machines. RESTful applica- tions use HTTP requests to post data (create and/or update), make queries, and delete data. Thus, REST uses HTTP for all four Create/Read/Update/Delete operations. REST is not a "standard" form. There will never be a World Wide Web Consortium (W3C) recommendation for REST, for example. And while there are REST programming frameworks, work- ing with REST is so simple that you can often come up with standard library features in languages like Perl, Java, or C#. OAuth is short for Open Authorization. OAuth provides a way for you to authorize an application to access data you have stored away in another application without having to share your username and password. You (the end user) want

to authorize an application of some sort (the client) to access some of your data (a scope) thats managed by a web service (the resource owner). Instead of asking for your password, the client redirects you to the resource owner, and you authorize a scope for the client directly with the resource owner. Assuming the end user authorizes the client, the client is notified and given an authorization code confirming that the end user has authorized it to access a scope. The client presents the authorization code it just received along with its client identifier and corresponding client secret to the resource owner and gets back an access token. The combination of client identifier, client secret, and authorization code ensures that the resource owner can positively identify the client and its authorization. The client uses the access token to make requests on behalf of the end user until the access token is revoked or expires.

3.3 REST API vs Streaming API:

The set of streaming APIs offered by Twitter give developers low latency access to Twitters global stream of Tweet data. A proper implementation of a streaming client will be pushed messages indicating Tweets and other events have occurred.4 Twitter offers several streaming endpoints, each customized to certain use cases. Public Streams- Streams of the public data flowing through Twitter. Suitable for following specific users or topics, and data mining. User Streams- Single-user streams, containing roughly all of the data corresponding with a single users view of Twitter. Site Streams- The multi-user version of user streams. Site streams are intended for servers which must connect to Twitter on behalf of many users.

3.4 Unix Shell Script:

A Unix shell is a command-line interpreter or shell that provides a traditional Unix-like command line user interface. Users direct the operation of the computer by entering commands as text for a command line interpreter to execute, or by creating text scripts of one or more such commands. Users typically interact with a Unix shell using a terminal emulator, however, direct operation via serial hardware connections, or networking session, are common for server systems. The shell definition line tells the system what program (shell) should be used to interpret the script's commands, and where the program (shell) is located. For the example shell script above, the shell definition line tells the system to use the Korn Shell (#!/bin/ksh) when executing this script.

For our purpose we have separated screen names from csv using Unix Shell Script. Code Snippet:

```
awk -F'@"print"@"2'tweets2F4new.csv|cut-d":"-s-f1|awk'if(length(1)-1 length(1)<=16)print'|cut-d""-f1|sort|uniq>usernames.csv
```

3.5 Tools Needed:

The set of streaming APIs offered by Twitter give developers low latency access to Twitters global stream of Tweet data. A proper implementation of a streaming client will be pushed messages indicating Tweets and other events have occurred.4 Twitter offers several streaming endpoints, each customized to certain use cases. Public Streams- Streams of the public data flowing through Twitter. Suitable for following specific users or topics, and data mining. User Streams- Single-user streams, containing roughly all of the data corresponding with a single users view of Twitter. Site Streams- The multi-user version of user streams. Site streams are intended for servers which must connect to Twitter on behalf of many users. We need to get Access tokens which are the permissions given by Twitter to access their API.

3.6 Proposed Solution:

The platform will allow people to indicate which issues they find most important in order to receive information about what purchases to make and actions to take for helping political and social causes. Our platform will use social media data such as Twitter, as well as public information from congress to make suggestions to citizens. With the help of twitter bot we can crowd source citizens to take micro-actions whether to support or oppose a political candidate or social cause. Twitter bot exert citizens to perform the micro actions of their general interests to help social and political causes.

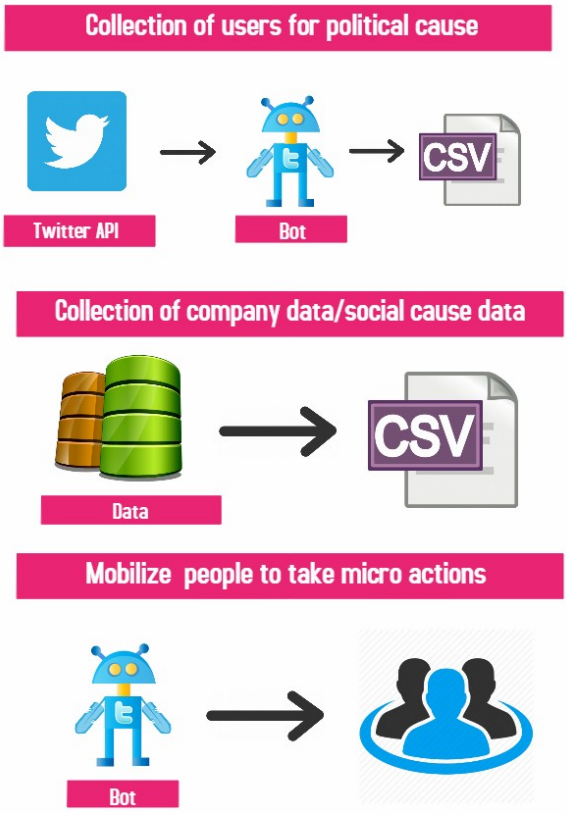

Figure 1. Flow of Data

3.7 Twitter Bot Configuration[9]:

A Twitter bot is a bot program used to produce automated posts on the Twitter micro blogging service, or to automatically follow Twitter users.

In order to code a bot first we need to create an account for bot in Twitter. Once the account is created create an application from developer interface of Twitter (dev.twitter.com). By following the terms and conditions from Twitter we should fill all the required details for the bot and generate Consumer Key, Consumer Key Secret, Access Token, Access Token Secret. In Application Permission model assign all the required permissions for the bot to perform required operations. The different types of permissions are Read Only, Read and Write, Read Write and Access Direct Messages. For this application required settings are Read and Write.

3.8 Python Script:

Python is a widely used high-level, general-purpose, interpreted, dynamic programming language. Its design philosophy emphasizes code readability, and its syntax allows programmers to express concepts in fewer lines of code than possible in languages such as C++ or Java. The language provides constructs intended to enable writing clear programs on both a small and large scale.

Python supports multiple programming paradigms, including object-oriented, imperative and functional programming or procedural styles. It features a dynamic type system and automatic memory management and has a large and comprehensive standard library.

3.9 Hashtags:

A hashtag is a type of label or metadata tag used on social network and microblogging services which makes it easier for users to find messages with a specific theme or content. Users create and use hashtags by placing the hash character (or pound sign #) in front of a word or unspaced phrase, either in the main text of a message or at the end. Searching for that hashtag will yield each message that has been tagged with it.

3.10 Data Collection and Refinement:

With the help of Python Script using any hashtag collect the tweets by crawling on stream data in Twitter.

Python Script for this functionality:

```
twitter_stream = Stream(auth, listener())
twitter_stream.filter(track ='#trumpyourthanksgiving')
```

Once all the tweets are collected they will be exported into a csv file. This csv file is further refined using Unix shell script to get the list of names tweeted with that hashtag.

3.11 Statistics of Data collected with time:

Number of dynamic stream tweets collected in an hour shows direct proportionality with time. This drastic increase of tweets leads to very huge amount of data. This big data is handled by running timer jobs on it.

3.12 Assigning Micro Actions to crowd:

To the citizens micro actions are suggested by the bot by tweeting them personally for social or political cause[10].

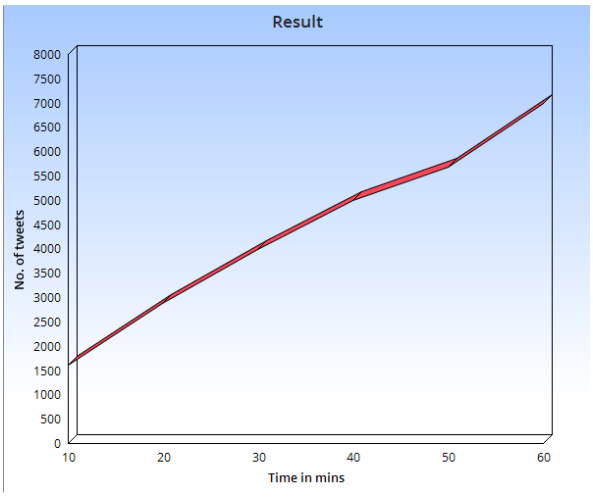

Figure 2. Tweets retrieved in One hour

Python Script for this functionality:

```
for i in range(0, len(a)):
message = "Show your thankfulness to Trump by paying through Paypal!! %s" %a[i]
api.update_status(status = message)
```

4. RESULT AND ANALYSIS

From Figure 2. we can understand that 7398 tweets retrieved for the hashtag trumpyourthanksgiving during Thanksgiving time,2016 in a span of 60 minutes. All these tweets are further refined, to get the list of users who supports Trump and interested in Thanksgiving shopping. For our experiment we have suggested these users with a suggestion Show your thankfulness to Trump by paying through Paypal!!. As Paypal supports Trump, we encouraged all citizens to pay through Paypal which raises funds to Paypal and helps Trump. In this way we succeeded implementing our proposed system.

5. CONCLUSION

In this paper, we have collected tweets from twitter using some political and social hashtags. We refined the data and collected user screen names. We have sent general messages and suggestions related to the hashtags to the groups of people who have tweeted on that hashtag. We can send many suggestions to the citizens to take actions on political and social causes by using this bot. The suggestions given by the twitter bot made citizens to take micro actions on that political and social causes. These suggestions are not only limited to twitter we can create bots in different social media platforms like Facebook, LinkedIn etc. and we can send suggestions to the people.

6. LIMITATIONS

1. As of now we have worked on streaming data from the twitter but not on the past tweets. 2. With the help of this twitter bot we cannot send messages to large number of users at a time.

7. FUTURE WORK

We will extend our twitter data collection by using specific words and images extension to hashtags. As of now we gave general suggestions by using a hashtag to the users. In fu- ture we will give matching suggestions to the user interest that help political and social causes. Giving matching suggestions involves big data analytics and some of the matching algorithms. We will work on the limitations mentioned in the paper.

8. ACKNOWLEDGMENTS

We would like to express our deepest appreciation to **Dr.Saiph Savage** who continuously motivated us to achieve success in each step. A special gratitude to **Dr. Donald Adjeroh**, whose contribution in stimulating suggestions and encouragement, helped us to coordinate our project.

9. REFERENCES

[1] A. Samuel, How Bots Took Over Twitter, Harvard Business Review, 19-Jun-2015. [Online]. Available: https://hbr.org/2015/06/how-bots-took-overtwitter.

[2] F. Morstatter, J. Pfeffer, H. Liu, and K. M. Carley, Is the Sample Good Enough? Comparing Data from Twit- ters Streaming API with Twitters Firehose, arXiv:1306.5204 [physics], Jun. 2013.

[3] Z. Chu, S. Gianvecchio, H. Wang, and S. Jajodia, Who is tweeting on Twitter: human, bot, or cyborg?, in Proceedings of the 26th annual computer security applications conference, 2010, pp. 2130.

[4] Cook, David, Waugh, Benjamin, Abdinpanah, Maldini, Hashimi, Omid, and Rahman, Shaquille Abdul, Twitter Deception and Influence: Issues of Identity, Slacktivism, and Puppetry, Journal of Information Warfare, vol. 13, no. 1.

[5] P. N. Howard, Pax Technica: How the Internet of Things May Set Us Free. New Haven, CT: Yale University Press, 2015.

[6] D. W. Butrymowicz, Loophole.com: How the Fecs Failure to Fully Regulate the Internet Undermines Campaign Finance Law, Columbia Law Review, pp. 17081751, 2009.

[7] P. N. Howard, New Media Campaigns and the Managed Citizen. New York, NY: Cambridge University Press, 2006.

[8] M. C. Forelle, P. N. Howard, A. MonroyHernandez,and S. Savage, Political Bots and the Manipulation of Public Opinion in Venezuela, Project on Computational Propaganda, Oxford, UK, Working Paper 2015.1, Jul. 2015.

[9] P. N. Howard and B. Kollanyi, Bots, StrongerIn, and Brexit: Computational Propaganda during the UK-EU Referendum, arXiv:1606.06356 [physics], Jun. 2016.

[10]TwitterbotConfiguration: https://github.com/codeforboston/CutePetsBoston/wiki/Setting-Up-a-Twitter-Bot