Identify The Rate Of Shares Of Each Fake News Source From Twitter

CS232 Big Data Problem Using MongoDB*

Kier S. Guevara
Ateneo de Manila University
Quezon City, Philippines
kier.guevara@gmail.com

Aris Gail S. Mendoza
Ateneo de Manila University
Quezon City, Philippines
aris.mendoza@obf.ateneo.edu

ABSTRACT

This paper provides a demonstration of using MongoDB in solving a Big Data problem wherein the author will select all tweets pertaining to Philippine Government with links to fake news sources. Using MapReduce feature, the authors will show how many articles were shared per fake news source from the data source, as well as how many tweets and how much retweets were done for those tweets.

This paper will also contains the steps how the authors had done the replication and sharding of the database. In the end, they will conclude if MongoDB was able to provide enough tools to show the required outputs.

ACM Reference Format:
Kier S. Guevara and Aris Gail S. Mendoza. 2017. Identify The Rate Of Shares Of Each Fake News Source From Twitter: CS232 Big Data Problem Using MongoDB. In *Proceedings of CS232 Project Defense (CS2322017).* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Right now, fake news has become widespread in the country, that if a person wouldn't be careful and curious enough, they would fall for the wrong information. These articles are widespread on the Internet, moreover in the social media like Facebook and Twitter. Good thing several institutions had provided lists of all fake news sources that people can refer to once they encounter articles found doubtful to believe. These lists will be helpful in identifying those dubious articles if it is fake or not.

1.1 Objectives

The authors aimed to do the following on this project:

(1) Identify a big data problem to solve using MongoDB
(2) Enumerate the steps in accomplishing the big data problem
(3) Visualize the results of the solution
(4) Perform replication and sharding operations in MongoDB

1.2 Research Questions

This project will answer the following questions:

(1) How many tweets contains links to different websites?
(2) What are the top websites do tweets contain from?
(3) How many tweets contain links to fake news sources?
(4) Which keywords are mostly found in tweets containing links with fake news sources?
(5) What is the rate of fake news shared against the total tweets pulled?

1.3 Scope and Limitation

This project will concentrate on the following scenarios to answer the research questions above:

(1) Only articles shared via Twitter are used in this study.
(2) Extracted tweets were based on keywords pertaining to Philippine Government from December 3, 2017 and 7 days backwards.
(3) List of fake news sources is based on the released list of CBCP appeared in GMA News Online website last June 27, 2017.
(4) Links coming from URL Shortening services(e.g., bit.ly, goo.gl, etc.) may not be included in this project.

2 REVIEW OF RELATED LITERATURE

2.1 The Fake News Source List

The Catholic Bishops Conference of the Philippines had released a list of websites spreading fake news last June 22, 2017 weeks after Justice Secretary Vitaliano Aguirre alleged during a press conference last June 7 that opposition party lawmakers of the Duterte administration were involved in clashes between government security forces and the Islamic State-inspired Maute Group in Marawi City[2][1] . But this was proven wrong after the photo presented by the Justice Secretary was discovered to be taken from its owner's album way back 2015. But then he told that he was 'misquoted' by the media, as this was spread all over social media as well, then after apologized. This incident had triggered CBCP to take measure and released the aforementioned list, saying spreading fake news is a sin. The list, which was partial of that moment, are the following[3][2]:

- Pinoy Trending - http://pinoytrending.altervista.org/
- Pinoy Trending News - http://pinoytrendingnews.net/
- Public Trending - http://www.publictrending.net/ (⇓)
- FilipiNews PH - http://www.filipinewsph.com/ (⇓)
- Trending News Portal - http://www.trendingnewsportal.net.ph/ (*Under Construction*)
- Classified trends - http://www.classifiedtrends.net/ (⇓)

*A partial fulfillment of the course requirement for CS232 Database Systems

CS2322017, December 2017, AdMU, Quezon City, PH
2017. ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

[1] http://cnnphilippines.com/news/2017/06/22/cbcp-fight-fake-news.html
[2] http://www.gmanetwork.com/news/news/nation/616023/cbcp-releases-list-of-fake-news-sites/story/

- Definitely Filipino - http://definitelyfilipino.com/
- Duterte News Info - http://www.du30newsinfo.com/ (*Under Construction*)
- Extreme Readers - http://www.extremereaders.com/ (⇓)
- Get Real Philippines - http://www.getrealphilippines.com/
- Guard1an - https://theguard1an.com/ (⇓)
- Kalye Pinoy - http://www.kalyepinoy.com/ (*Redirected to an extension install page, possible virus*)
- Leak News PH - http://www.leaknewsph.com/ (⇓)
- Media ni Duterte - http://www.dutertedefender.com/
- Minda Nation - http://mindanation.com/
- Netizens PH - http://www.netizensph.com/ (*Redirected to a different page, possible virus*)
- News Media PH - http://www.newsmediaph.com/
- News Titans - http://www.newstitans.com/
- OKD2 - http://okd2.com/ (http://www.okd2.com)
- Pinoy Freedom Wall - http://www.pinoyfreedomwall.com/ (⇓)
- Pinoy Viral Issues - http://www.pinoyviralissues.net/ (*Under Construction*)
- Pinoy Viral News - http://pinoyviralnews.com/
- PinoyWorld - http://www.pinoyworld.net/ (*Redirected to a different URL, but* ⇓)
- SocialNewsPH - http://www.socialnewsph.com/ (*Selected readers only*)
- TahoNews - http://www.tahonews.com/ (⇓)
- TheVolatilian - http://www.thevolatilian.com/
- Thinking Pinoy - http://www.thinkingpinoy.net/
- TrendingBalita - http://www.trendingbalita.info/ (⇓)
- TrendTitan - http://trendtitan.com/ (*Under Construction*)

2.2 Social Media As The Driver To Spread Fake News

Using social media networks like Facebook, Twitter, and Instagram made the world even smaller due to the fact that it can connect people anywhere. Thanks to the Internet it has become possible. Integrating it in daily lives had made such a significant change, to the point that most of people rely on it in every aspect of life - house and living tips, cooking, health and lifestyle, business and e-commerce, entertainment and games, and even getting updates from news. Based on the cited research[6][3] by BBC on their article last November 6, 2016, "...suggesting an increasing proportion of US adults are getting their news from social media..."[7][4] and another one cited from Adornato[1][5] that a third of the news bulletins of the TV stations in US had reported information coming from social media, indeed that social media is a potential medium to reach every person in the Earth.

With the power of social media to transmit information anywhere, thus a new issue arises: not all information posted on them are true and accurate. It may not hurt to post those kind of articles on the web. But if these articles were continuously spread all over the internet and many people believe in what it says and persons involved are alarmed and are the subject of conversations, then it is a different story. Most fake news stories, quite sadly to say, were even shared by famous people, some with incredible credentials, and mostly with a big network of followers, which makes these stories even convincing. These people, with a huge network of followers and have the tendency to share anything just to continue gain popularity, are those targeted to be the catalyst to spread it. They were selected based on the data being sold by tech companies to advertisers. These data, if not used correctly, will be used by the purveyor to get these groups of people and seed information to them, on which, in return, will comment or share or recommend it on their social networks. This will spread more when it reaches those poorly informed people who had hard time claiming these information if it is true or not[4][6].

Education may be one of the solutions to stop spreading fake news. While people know that fake news exists and their purveyors are active, giving them the chance to learn strategies in identifying them will further lessen the chance these fake news will spread. Given this knowledge, social networks like Facebook and Google are giving people the chance to help them stop the spreading of fake news by tagging and reporting these articles as fake or inappropriate or spam so that these articles won't circulate in their network. These efforts of these social networks help the society in promoting true and accurate information on the Internet.

3 METHODOLOGY

3.1 Pull Data from Twitter and Clean It

3.1.1 Pull Data from Twitter. The authors decided to use a Python library named **`twarc`** to pull tweets from Twitter. Their program **`fakeNews.py`**, written in Python using this library, is an efficient and enough tool to pull specific data based on a search criteria. It requires internet connection to run. The search criteria are as follows:

- The tweet is about Philippine Government. The authors used the following keywords in getting Tweets from Twitter:
  - *Duterte*
  - *Leila De Lima*
  - *Philippine Congress*
  - *Philippine Corruption*
  - *Philippine Crimes*
  - *Philippine Election*
  - *Philippine Government*
  - *Philippine Judiciary*
  - *Philippine Legislative*
  - *Philippine Scandal*
  - *Philippine Senate*
  - *Philippines Marcos*
  - *PNoy*
  - *PNP*
  - *Robredo*
  - *Trillanes*
- The tweet must contain links to certain websites. These links will be used to compare against the list of fake news source in the coming steps.

[3] http://www.journalism.org/2016/05/26/news-use-across-social-media-platforms-2016/

[4] http://www.bbc.com/news/blogs-trending-37846860

[5] http://anthonyadornato.com/study-reveals-social-medias-impact-on-editorial-decisions/

[6] https://www.bloomberg.com/view/articles/2017-08-31/why-fake-news-spreads-so-fast-on-facebook

To run the program, the computer to be used must have the following software installed:

- Python 3.6.3 (Windows/MacOS/Linux, Ubuntu have version 2.7.3 installed by default)
- The following Python libraries, installed using `pip` command:
  - twarc
  - request
  - requests

On the command line, make sure the prompt is directly where the Python file resides, then do the following:

(1) Open the `fakeNews.py` using your preferred text editor.
(2) Inside the file go to `Line 135`, you will see the following code:

```
strQuery = "Philippines filter:links"
```

Leave the `filter:links` words intact and just replace the first word inside the quotation marks with the search key to search in Twitter. Make sure there is a space in between the search key and the `filter:links`.
(3) Save changes on the file and close it.
(4) On the same prompt in the terminal, issue the command:

```
python fakeNews.py
```

to run the program. It will halt once it finished processed all the retrieved Tweets or if the records reached 1000000.

Note: While running the app, there may times that the processing stops in some instances, but the internet connection is still up. To resolve this, just press Enter key or Arrow Down key to make it continue, or just make sure the screen wouldn't sleep during processing.

3.1.2 Clean Up Data. Once data has been retrieved from Twitter, the program will further clean it up by filtering Tweets with URLs coming from URL shortening services that was extracted from the `extended_url` field. The program will try to open the link and get the actual URL of it. If it is successful, it will be saved in `fileCSV_fullUrl.csv`, and those which have failed to retrieve will be saved in `fileCSV_shortUrl.csv`. Furthermore, those Tweets with valid URLs will be automatically saved in `fileCSV_fullUrl.csv`.

These text files will be CSV-formatted. Please see the **Table 1** for the details.

3.2 Replicate and Shard

Preparing the MongoDB servers is essential in keeping data in the database. Making it more accessible and reliable will make it failsafe. Next steps will demonstrate how to create replica sets and shard servers in the database system. This activity will enable the data to be distributed the data on different instances of MongoDB to maximize the write and read operations, as well as availability of data will be reliable if in any case a server instance is down.
All the steps done here are based on the documentation as prescribed by MongoDB[5].

Note: If planning to use separate workstations, make sure the installation of MongoDB on all workstations have the same version number; for release candidates, make sure all of them uses the same release candidate version. Likewise, configure each workstations' firewall settings to either allow the ports specified below to pass via `UDP` *protocol, allow the MongoDB executable files (*`mongod, mongo, mongos, mongoimport,` *etc), or allow specific IP addresses so they can communicate with each other.*

3.2.1 Replicate. For each item, follow the configuration file format at indicated in the *Appendix A*.

(1) Config Servers
Replica Set: `CS232-cfg`
Ports:
- 27018
- 27019
- 27020

Command to run the instance of the server: (Do this for each port)

```
mongod --config <path to your config file>
```

Connect to one of the server instance using the following command:

```
mongo --host <host, ip, or localhost> --port <assigned port>
```

On the `mongo>` prompt, initialize the config servers by issuing this command:

```
mongo>rs.initialize(
... {
...   _id: "<replica set name>",
...   configsvr: true,
...   members: [
...    { _id: 0, host: "<host, ip, or localhost>:<portNo>"},
...    ...,
...    { _id: N, host: "<host, ip, or localhost>:<portNo>"}
...   ]
... }
...)
```

(2) Shard Servers
- Replica Set: `CS232-shr1`
  Ports:
  - 27021
  - 27022
  - 27023
- Replica Set: `CS232-shr2`
  Ports:
  - 27025
  - 27026
  - 27027
- Replica Set: `CS232-shr3`
  Ports:
  - 27028
  - 27029
  - 27030

Command to run the instance of the server: (Do this for each port)

```
mongod --config <path to your config file>
```

For each replica set, connect to one of the server instance using the following command:

```
mongo --host <host, ip, or localhost> --port <assigned port>
```

On the `mongo>` prompt, initialize the shard servers by issuing this command:

Table 1: The fields included in the output CSV file

Fields	Data Type	Description
userID	Long	The ID number corresponds to the actual user
userScreenName	String	The user's handle in Twitter
userFollowersCount	Integer	The number of followers the current user have
userFriendsCount	Integer	The number of users the current user follows
userFavoritesCount	Integer	The number of tweets the current user had liked
tweetID	Long	The ID number corresponds to the current tweet
dateCreated	Date	The date when the current tweet is created
fullText	String	The text contained in the current tweet
userLocation	String	The friendly name of the location where the tweet came from
expandedURL	String	The original URL where the Twitter URL shortening service link points to
mainURL	String	The domain of the expandedURL
retweetCount	Integer	The number of reshares or retweets the current tweet gets
favoriteCount	Integer	The number of likes or favorites the current tweet gets
isRetweet	String	A flag to determine if the current tweet is the original tweet or just a retweet
originalTweetID	Long	The original tweet where this retweet came from
searchKeyword	String	The search keyword used in `fakeNews.py` program to pull the current tweet

```
mongo>rs.initialize(
... {
...   _id: "<replica set name>",
...   members: [
...     { _id: 0, host: "<host, ip, or localhost>:<portNo>"},
...     ...,
...     { _id: N, host: "<host, ip, or localhost>:<portNo>"}
...   ]
... }
.. )
```

Note: Take note which shard instance was set to `PRIMARY`*. This information will be used in adding shards to the* `mongos` *instance later.*

Given all the steps have been executed properly, replication has been completed. Go to next step to perform sharding.

3.2.2 *Shard.*

(1) Configure the Mongo Router. Create a config file based on the `Router` entry in `Appendix A`. Use the port number `27024` and save it in the config file under:

```
net:
  port: 27024
```

(2) Make sure to specify the config servers correctly as the Router will base its configuration from them.
(3) Save the configuration file.
(4) Run the Mongo Router on the terminal by issuing the command:

```
mongos --config <path to config file>
```

(5) Open another terminal and connect to the `mongos` instance using the command:

```
mongo --host <host, ip, or localhost> --port <portNo>
```

Port number is the port number assigned in the configuration file of the `mongos` instance.

(6) On the `mongos>` prompt, add one shard server from each shard replica set using the command:

```
mongos> sh.addShard(
        "<replicaSetName>/<host, ip, or localhost>:<portNo>")
```

Note: Add the `PRIMARY` *server for each shard replica set.*

(7) Create a database on the `mongos` instance and a collection inside the created database using any method (GUI or terminal).
(8) To enable sharding on the created database, issue the command:

```
mongos> sh.enableSharding("<databaseName>")
```

(9) Use the created database and create a hash index of `_id` on it. Use the command:

```
mongos> use <databaseName>
mongos> db.collection.createIndex({<key>: "hashed"})
```

This index will serve as the Router's basis in distributing the data over the shard replicas.

(10) Once index has been created, tell `mongos` that the data in the selected table that it will be distributed among the shard sets. Use the command:

```
mongos> sh.shardCollection("<database>.<collection>"
        , {"_id": "hashed"}, false
        , {numInitialChunks: 5})
```

After accomplishing the above steps, the database is ready to accept data to its target collection. Once data came in to that collection, the `mongos` will automatically route the data to the appropriate shard replica.

3.3 Upload Data to MongoDB

The cleaned data from the text file `fileCSV_fullUrl.csv` will be uploaded either by:

- GUI-based IDE for managing MongoDB (*Studio 3T*)
- terminal command

```
mongoimport --host <host, ip/ or localhost> --db <database>
        --collection <collection> --type csv
        --headerline --file <path to the CSV file>
```

[5] This method was used in uploading the data to the MongoDB[5].

3.4 Perform MapReduce on Data

MapReduce[5] function is similar to SQL's Group By or Aggregate Functions that will summarize data based on a criteria. The authors will perform MapReduce functions to achieve the following data:

- Get the total tweets for each URL

```
//get the total count of tweets for each main url
var mapMainURL= function(){
        emit(this.mainURL,1);
}

//get the total count of distinct main
//url from the tweets without the
//searchKeyword:Philippines
var mapMainURLWOPh = function(){
        if(this.searchKeyword != "Philippines")
                emit(this.mainURL,1);
}

var reduceArraySum = function(key,value){
        return Array.sum(value)
}

db.fakeNews.mapReduce(mapMainURLWOPh,
        reduceArraySum,{out:"totalMainURLWOPh"})
```

- The rate of shares of each fake news source from Twitter

```
var mapFakeNewsSite = function(){
                emit(this._id.mainUrl ,{count:this.value});
}

var reduceArraySum = function(key,value){
        var count = 0;
                for(var i = 0; i < value.length; i++)
                        count += value[i].count;
        return {count:count};
}

db.fakeNewsSiteWithKeyword.mapReduce(
        mapFakeNewsSite,reduceArraySum,{out:"fakeNewsSite"})
```

- Keywords mostly found in tweets containing links of fake news sources

```
var isFakeNews = function(url){
        var listFakeNews = ["http://www.classifiedtrends.net/",
                "http://definitelyfilipino.com/",
                "http://www.extremereaders.com/",
                "http://www.filipinewsph.com/",
                "http://www.getrealphilippines.com/",
                "https://theguard1an.com/",
                "http://www.kalyepinoy.com/",
                "http://www.leaknewsph.com/",
                "http://www.dutertedefender.com/",
                "http://mindanation.com/",
                "http://www.netizensph.com/",
                "http://www.newsmediaph.com/",
                "http://www.newstitans.com/",
                "http://okd2.com/",
                "http://www.pinoyfreedomwall.com/",
                "http://pinoytrendingnews.net/",
                "http://www.pinoyviralissues.net/",
                "http://pinoyviralnews.com/",
                "http://www.pinoyworld.net/",
                "http://www.tahonews.com/",
                "http://www.thevolatilian.com/",
                "http://www.socialnewsph.com/",
                "http://www.publictrending.net/",
                "http://pinoytrending.altervista.org/",
                "http://www.du30newsinfo.com/",
                "http://www.thinkingpinoy.net/",
                "http://www.trendingbalita.info/",
                "http://trendtitan.com/",
                "http://www.trendingnewsportal.net.ph/"
        ]

        for(var i = 0; i < listFakeNews.length; i++)
                if(listFakeNews[i] == url)
                        return 1;

        return 0;
}

db.system.js.save( { _id:"isFakeNews",value:isFakeNews})

var mapFakeNewsSiteWithKeyWord = function(){
        if(isFakeNews(this.mainURL) > 0)
                emit({mainUrl:this.mainURL,
                searchKeyword:this.searchKeyword} ,1);
}

var reduceArraySum = function(key,value){
        return Array.sum(value)
}

db.fakeNews.mapReduce(mapFakeNewsSiteWithKeyWord,
        reduceArraySum,{out:"fakeNewsSiteWithKeyword"})
```

To perform the above MapReduce operations, the user should be connected to the `mongos` server. See the command below:

```
mongo --host <host, ip, or localhost> --port <portNo>
```

The port number for the `mongos` server is indicated in the *Shard* subsection.

Once connected, the `mongos>` prompt will appear. On the prompt, type the MapReduce function as stated on the list above.

Output data of each MapReduce activity will be saved in their respective collections.

3.5 Data Visualization

The authors used Microsoft Excel as a means of visualizing the data gathered from the MapReduce operations. MS Excel had all the required tools in creating visualizations like Charts. For them to do this, data gathered from MapReduce will be exported using the following command:

```
mongoexport --host=<host, ip, or localhost
        --db <database> --collection totalMainURLWOPh
        --type=csv --fields _id,value
        --out <path where to save>
```

[5] The output will be a CSV file that will be imported to MS Excel where the visualization will be made.

4 RESULTS AND ANALYSIS

After all those queries on the imported data with more than 30,000 records, here are the notable results that were found by the authors:

(1) On `Figure 1`, the graph shows the total number of tweets pertaining to each website. This is based on the data that was imported to the MongoDB setup made by the authors. It shows the distribution of tweets on each website, telling us where the tweets' subject came from. Most of the tweets containing links that comes from social media networks like Twitter and Facebook. Also, it is worth noting that articles published from the website Rappler are quite abundant. We can also see that people from Twitter loves sharing videos from YouTube.

Aside from those website, people using Twitter are well informed when it comes to news and current affairs as many tweets were articles published from notable local and international news channels.

Using the same data, the authors had collected the top 10 websites which the tweets are about (`Figure 2`).

Collecting all the top 10 websites garnering the most tweets, the authors compare it against the tweets garnered by websites not included in the top 10. The tweets about the top 10 websites are overwhelmingly exceeds the others not in the top 10 (`Figure 3`).

(2) The authors then selected only the tweets that contains links from fake news websites. They compared the `mainURL` field of the raw data against the list of fake news websites as prescribed by the CBCP. From this comparison, here are the rate of tweets per fake news sources (`Figure 4`).
From the data gathered, the authors discovered that only 4 of the websites listed in the fake news sources have garnered tweets. These are:

(a) http://www.getrealphilippines.com - 97.42%
(b) http://pinoytrendingnews.net - 1.29%
(c) http://www.dutertedefender.com - 0.86%
(d) http://www.thinkingpinoy.net - 0.43%

(3) Based on the data from above, the authors also checked what keywords on those tweets were used. `Figure 5` shows the keywords distribution among those tweets containing links to fake news sources. The top 3 keywords are:

(a) `Duterte` - 48%
(b) `Trillanes` - 34%
(c) `Leni Robredo` - 11%

Using the number of tweets containing the links to fake news sources, the authors compared it against all other tweets gathered for this project (`Figure 6`). The tweets containing links to fake news sources are only 0.49% among all tweets they gathered before performing MapReduce.

5 CONCLUSION

From the retrieved tweets on the mentioned date range, the rate of tweets containing links from fake news source which are tagged by CBCP are just half of 1% of the total number of tweets. On this rate, only 4 websites from the list published by CBCP were being actively shared over Twitter. Possible that users are more aware about fake news and avoids them, or the list CBCP had released outdated.

Also, based on the top 10 results of sources of information in Twitter, most of the links came from social networks like Facebook, Twitter, and YouTube, and the rest are from local news websites like ABS-CBN, GMA, Inquirer, Rappler, and the like.

On the other hand, the tools the authors had used to accomplish this project, especially the MongoDB, had made this all possible.

APPENDIX

A CONFIGURATION FILE FORMAT FOR EACH MONGODB SERVER

The following configuration file formats are used in implementing the replica and shard servers in the author's setup.

A.1 Configsvr

```
# Where and how to store data.
storage:
  dbPath: <path to the data folder of the server>
# network interfaces
net:
  port: <port number assigned>
  bindIp: 0.0.0.0 #same as --bind_ip_all option
replication:
  replSetName: <replication set name>
sharding:
  clusterRole: configsvr
```

A.2 Shardsvr

```
# Where and how to store data.
storage:
  dbPath: <path to the data folder of the server>
# network interfaces
net:
  port: <port number assigned>
  bindIp: 0.0.0.0 #same as --bind_ip_all option
replication:
  replSetName: <replication set name>
sharding:
  clusterRole: shardsvr
```

A.3 Router

```
# network interfaces
net:
  port: <port number assigned>
  bindIp: 0.0.0.0 #same as --bind_ip_all option
sharding:
  configDB: <replicaSetName>/<configsvr1>, <configsvr2>, \\
  ..., <configsvrN>
```

ACKNOWLEDGMENTS

The authors would like to thank everyone who contributed to the success of this paper - our classmates in CS232 especially to the MongoDB masters Migy, Abe, and Pat: thank you for the support and knowledge you shared with us in completing this paper. For our friends to cheer us up when we are down. For our special someone to inspire us every time. For our parents with their never ending support and guidance. To our prof, Sir JC, who had provided us the knowledge and skills in accomplishing this database project. And last, but not the least, the One Above with His unending mercy and compassion. Everyone, MARAMING SALAMAT!

REFERENCES

[1] Anthony Adornato. 2016. Study Reveals Social MediaâĂŹs Impact on Newsroom Editorial Decisions. (2016). http://anthonyadornato.com/study-reveals-social-medias-impact-on-editorial-decisions/
[2] VJ Bacungan. 2017. CBCP to public: Fight 'fake news'. (2017). http://cnnphilippines.com/news/2017/06/22/cbcp-fight-fake-news.html

[3] Marlly Rome Bondoc/RSJ/KVD. 2017. CBCP releases list of fake news sites. (2017). http://www.gmanetwork.com/news/news/nation/616023/cbcp-releases-list-of-fake-news-sites/story/
[4] Mark Buchanan. 2017. Why Fake News Spreads So Fast on Facebook. (2017). https://www.bloomberg.com/view/articles/2017-08-31/why-fake-news-spreads-so-fast-on-facebook
[5] MongoDB Documentation. 2017. The MongoDB 3.6 Manual. (2017). https://docs.mongodb.com/manual/
[6] Jeffrey Gottfried and Elisa Shearer. 2016. News Use Across Social Media Platforms 2016. (2016). http://www.journalism.org/2016/05/26/news-use-across-social-media-platforms-2016/
[7] BBC Trending. 2016. The rise and rise of fake news. (2016). http://www.bbc.com/news/blogs-trending-37846860

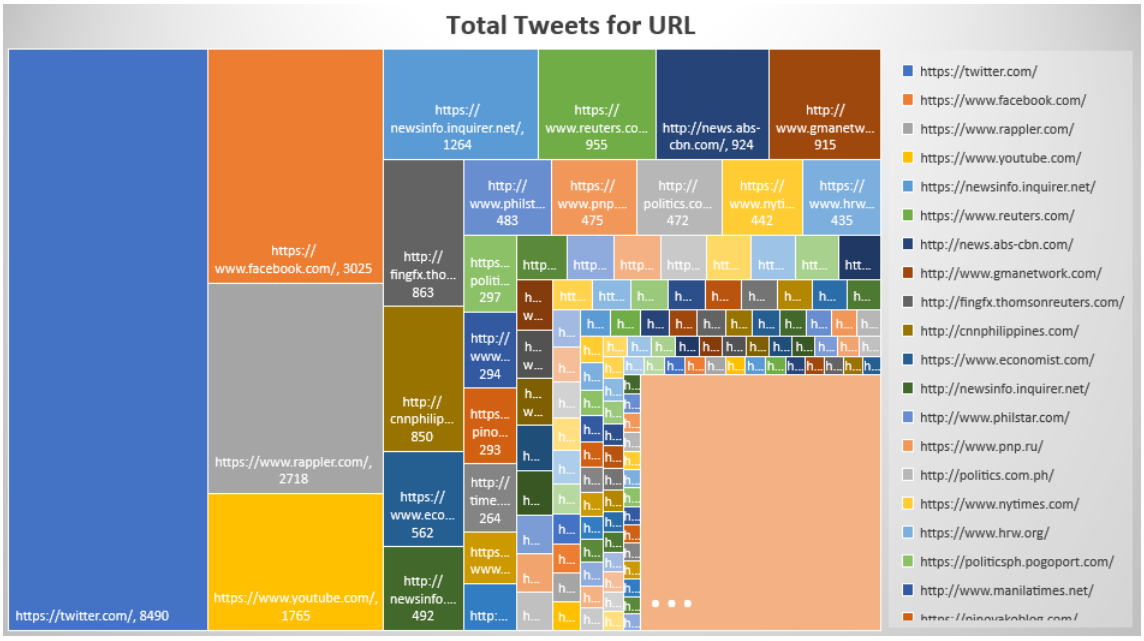


Figure 1: Total Tweets per URL

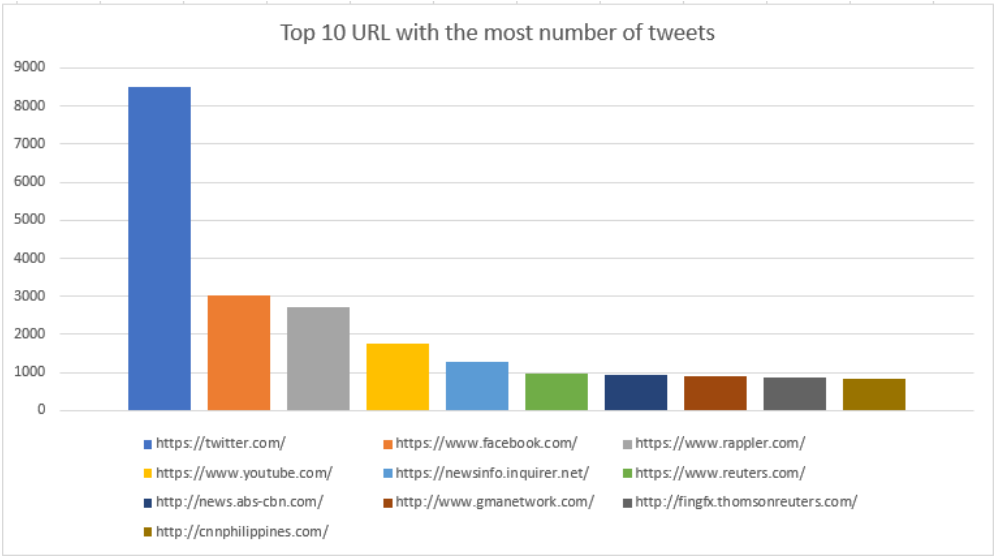


Figure 2: Top 10 URL with the most number of tweets

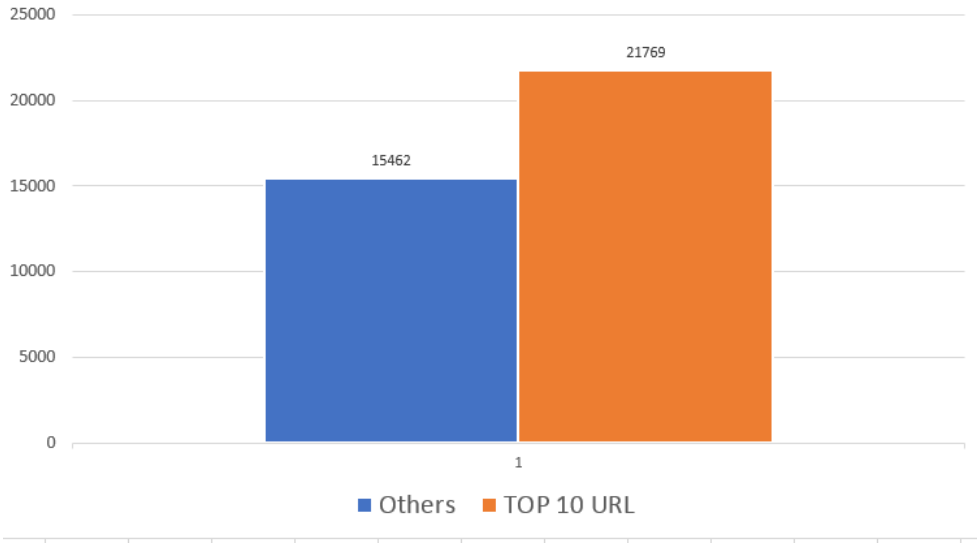


Figure 3: Comparison of number of tweets from the Top 10 websites against the others.

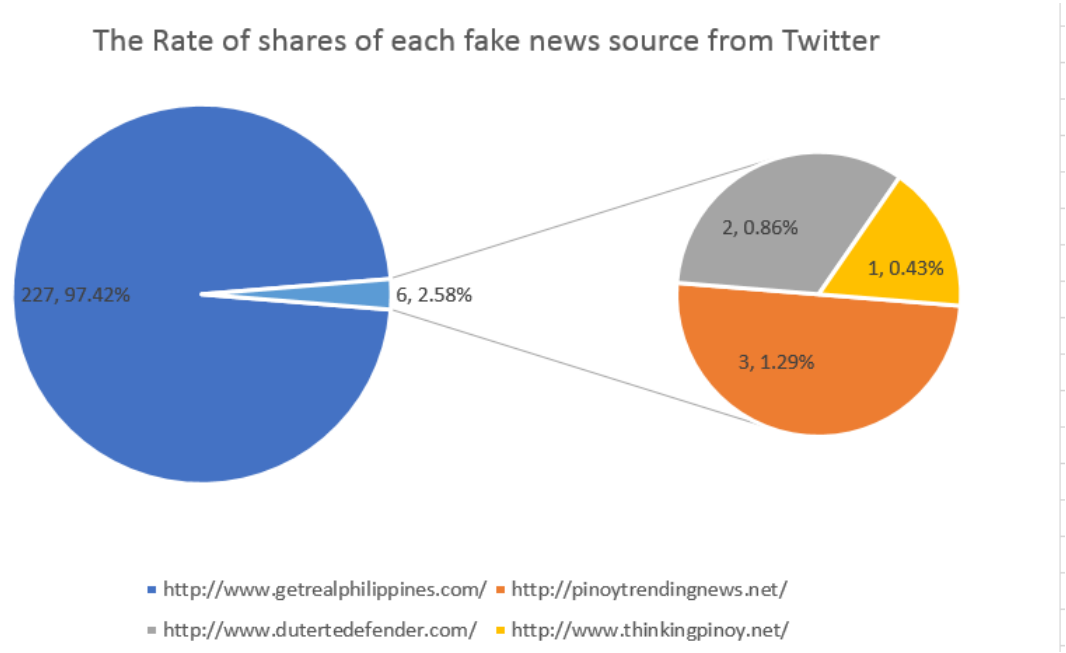


Figure 4: The Rate of Shares of Each Fake News Source From Twitter

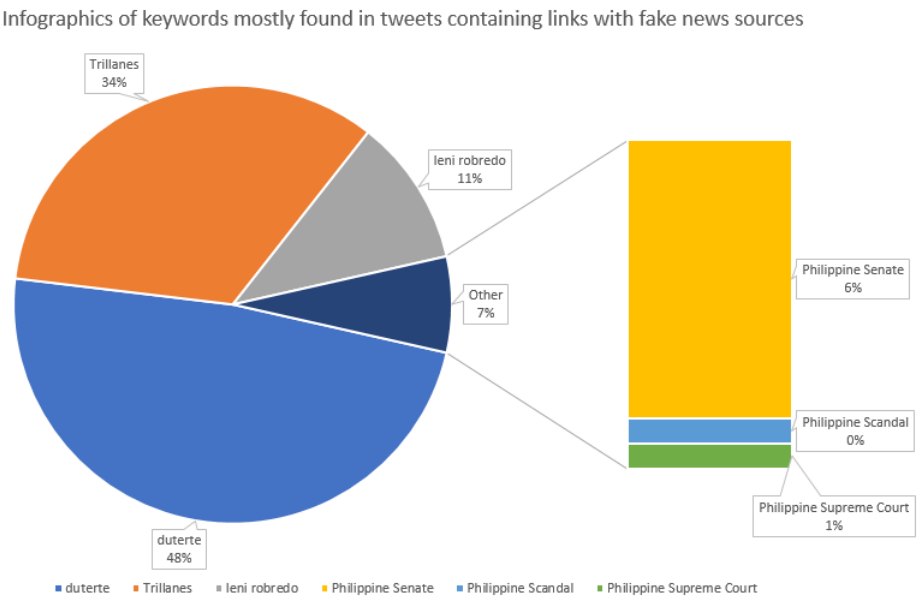
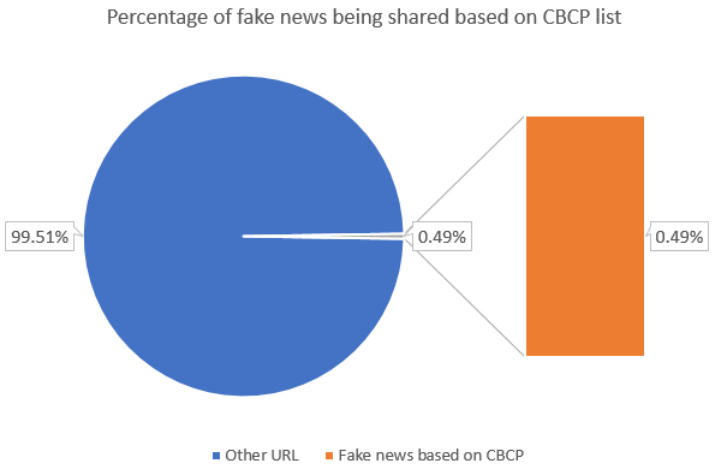


Figure 5: Infographics of Keywords Mostly Found In Tweets Containing Links To Fake News Sources



7.png

Figure 6: Percentage of tweets containing links to fake news against all other tweets.