Tracking Malaria Epidemic through Twitter

Aagam Shah, Michael Neuman, Yibing Shi

Department Of Electrical Engineering & Computer Science, Northwestern University
{aagamshah2017, MichaelNeuman2020, yibingshi2018}@u.northwestern.edu

***Abstract*—Social media has evolved as a platform for spreading information and opinions across the world. There has been a lot of research towards utilizing social media mining techniques to provide useful insights into medical research. Malaria is one of the many diseases which spreads rapidly and is mainly seen affecting developing countries and regions in Africa, South America and South Asia. Because of its prevalence, malaria causes hundreds of thousands of deaths each year. In this paper, we attempt to analyze tweets to determine the most discussed topics surrounding malaria as well as to identify initial indicators of an epidemic, medication, research on vaccines and their availability in the affected areas. We find that the most common topics discussed relating to malaria are often not useful in tracking malaria research and spread of the disease. We however find that areas with high volumes of tweets about malaria outbreaks have minimal overlap with areas tweeting about malaria vaccinations. This suggests that many regions that need vaccination and prevention techniques are not getting supplies quick enough.**

I. Introduction

MALARIA is a mosquito-borne disease that is caused by a parasite called *Plasmodium*. The parasite invades human red blood cells via interactions between the host and parasite surface proteins, causing fever, tired-feelings, vomiting, headaches and if left untreated, even seizures, coma, or death. Malaria is mainly transmitted from the bite of a female *Anopheles mosquito* [1].

In 2015, there were an estimated 212 million cases (around 2.9% of the world population) and 429,000 deaths as a result of the disease. A majority of the deaths were children in Africa. Around 1,500 cases of Malaria are diagnosed in the United States each year, the main people affected being travelers and immigrants coming from more malaria prevalent areas. In May 2017, The World Health Organization (WHO) announced a pilot program for the worlds first malaria vaccine which would first be implemented in three countries: Kenya, Ghana and Malawi [2]. There also has been much attention given to malaria by key donor institutions who wish to control and eliminate the disease. Organizations such as the Bill and Melinda Gates Foundation, Wellcome Trust, the Global Fund and the World Bank have provided millions of dollars in an effort to combat the disease [3].

Instructor: Professor Alok Choudhary, Electrical Engineering and Computer Science Department, Northwestern University.

The spread of malaria is most severe in developing African regions. With the rapid growth of social media networks, conversations about how to deal with the spread of the disease have become more open. This is because anyone with Internet access can communicate their ideas and creations across the globe. Social messaging services such as Twitter provide an instant stream of information being transmitted between people around the world [4]. Organizations, governments, and individuals can use this platform for broadcasting information on disease-related topics such as breaking news or statistics about global spreading of a disease like malaria.

If a pattern of malaria detection can be identified through observing social media platforms in real-time, this would benefit the research and study of malaria and assist in the effort to prevent and eliminate the disease. If a highly-concentrated outbreak can be detected early based on location, this could be used by organizations looking to quickly distribute prevention and treatment options. Although there are previous studies that focus on disease detection using social networks, most of them are limited to specific regions geographically (mainly the United States) and thus do not translate well to a disease like malaria which is most prevalent in Africa, south Asia and south America.

In this paper, we collected tweets containing keywords associated with malaria during April and May 2017. The focus of our study will be investigating and evaluating three representative weeks tweets. The remainder of this paper is organized as follows. We describe an overview of related work involving disease tracking through social media in Section 2. A description of our dataset and used tools follows in Section 3. In Section 4, we present our analysis on data we collected. Finally, in Section 5 we offer a few conclusions we have drawn as a result of our study as well as areas for future work.

II. Related Work

Some previous work exists involving utilization of a large stream of data in order to generate useful maps.

Middleton used statistical analysis from real-time tweet streams in order to generate a map of specific locations impacted by a natural disaster crisis. Looking at two specific case studies (Hurricane Sandy from October 2012 and Oklahoma

Tornado from May 2013), Middleton was able to generate a highly precise geo-parsing algorithm that utilized information from large location databases. During a crisis, minimizing false positives (maximizing precision) is of high priority so that emergency responders can focus on those users who are in the most at-risk areas. Middletons model has a map threshold which can be adjusted in order to obtain maximize precision [5].

Signorini investigated the relationship between Tweet volume and the spread of the H1N1 virus in the United States. They discovered that Twitter traffic can be used to track users interests and concerns associated with the H1N1 virus. They also were able to provide real-time estimates of Influenza-like illnesses, outpacing the traditional public health reports about disease activity which are normally posted up to 1-2 weeks after collecting data [6].

Missier investigated Twitter data to track Dengue outbreaks utilizing both LDA-based clustering and supervised clustering. They sorted Tweets containing Dengue-related keywords into four categories based on content (news, joke, mosquito focus, sickness). The authors noted that each clustering method had its drawbacks: LDA-based clustering doesnt work well went searching for pre-filtered terms and supervised classification requires manual annotation of tweets and may not be time-efficient [7].

Choudhary utilized spatial, temporal and text mining on Twitter data to design disease surveillance maps for the United States. These maps can be utilized for early prediction of seasonal disease outbreaks, like the flu, as well as monitoring the popularity of different treatments and cancer times based on location [8]

III. Experiment Setup

In this section, we describe how our data was collected and processed. In order to interact with Twitters streaming API, we used the python library Tweepy.

A. Dataset

We collected tweets during April and May 2017. After our preliminary findings, we designated three time periods to compare results between (further discussion in Section 4). Tweets from Week 1 are those collected from April 21-May 3 (defined as a special week in further discussion), Tweets from Week 2 are those collected on May 12-18, and Tweets from Week 3 are those posted during May 19-25.

B. Keywords

In order to encompass a large amount of data, we utilized a list of keywords related to malaria that we could search for on Twitter. The keywords spanned across multiple categories such as treatment and prevention, along with other useful words we found by conducting initial small sample experiments.

Our initial list of keywords used for our first few weeks of data collection included the following key words: *Malaria, Malaria vaccine / malaria vaccination, Malaria treatment, Malaria endemicity, Malaria blood sample, Malaria infection, Malaria economic, Malaria cost*. There are five types of Plasmodium parasites can infect humans for malaria [1]. Thus, we added the five parasite names *P. falciparum Malaria, P. vivax Malaria, P. ovale Malaria, P. malariae Malaria, P. knowlesi Malaria* to our initial keywords list.

After a brief period of collecting initial sample data, we evaluated our results and adjusted our keyword searches. Keywords such as *Malaria endemicity, Malaria blood sample, Malaria economic* and the five different types of malaria mentioned did not give us a high enough volume of tweets. When analyzing the set, we also discovered certain locations that occurred most often in our data set. We added these along with the other key phrases that produced a high enough volume of tweets. Our final list of key phrases that were searched on were:

malaria, malaria vaccine, malaria vaccination, malaria treatment, malaria outbreak, Malaria Kenya, Malaria Ghana, Malaria Malawi, Malaria Nigeria

In total we collected about 57,000 tweets containing one or more of the key phrases. For those three specific weeks, we have over 29,000 tweets which we analyzed.

C. Tools Used

The main tools used in the evaluation are (i) Scikit-learn[1], a machine learning algorithm for clustering; (ii) plot.ly[2], an online data analytics and visualization tool for graphing, analytics, and statistics; (iii) nominatim[3], a search engine for querying a name or address to get its geographic coordinate in json; (iv) geoplotlib[4], a python toolbox for visualizing geographical data and making maps; and (v) geoText[5] and (vi) geograpy[6], two packages that extract cities or countries from a given text or url. Below is a more detailed description of how some of these tools are used, specifically the case where we performing content clustering.

[1] http://scikit-learn.org/stable/
[2] https://plot.ly/
[3] https://nominatim.openstreetmap.org/
[4] https://github.com/andrea-cuttone/geoplotlib
[5] https://pypi.python.org/pypi/geotext
[6] footnote: https://pypi.python.org/pypi/geograpy

D. Content Clustering Method

In order to cluster the contents of the tweets, we needed an approach to help us perform text clustering.

Scikit-learn is a common simple and efficient machine learning tool used for data mining and data analysis [9]. Since our dataset is unlabeled, we used the clustering tool *sklearn.cluster* for categorizing tweets. The main method from this module we used here is *k-means* which partitions *n* observations into *k* clusters where each observation belongs to the cluster with the nearest mean [10].

k-means algorithm can be described as: given an initial set of *k* means $m_1^{(1)}, \ldots, m_1^{(k)}$, the algorithm proceeds by alternating between two steps:

- Assign each observation to the cluster whose mean yields the least within-cluster sum of squares (WCSS). Since the sum of squares is the squared Euclidean distance, this is intuitively the "nearest" mean:

$$\arg\min_s \sum_{i=1}^{k} \sum_{x \in S_i} \|x - \mu_i\|^2 = \arg\min_s \sum_{i=1}^{k} |S_i| \, Var S_i$$

  where μ_i is the mean of points in S_i. This is equivalent to minimizing the pairwise squared deviations of points in the same cluster.
- Calculate the new means to be the centroids of the observations in the new clusters:

$$\sum_{Cluster_i} \sum_{Dimension} \sum_{x,y \in C_i} (x_d - y_d)^2$$

We also design an algorithm to help us calculate the volume of tweets that could be classified to each cluster.

In Algorithm 1, we define tweet content as an array of words *T* and cluster from k-means as matrix *MC* which contains the keywords of each cluster. We begin by checking each word in *T* with each keyword in a cluster in *MC* and use an array *Point* to count the number of occurrences of a word that match with the keywords in each cluster. We then find the most matched cluster for this tweet by finding the maximum value of the *Point*. If the tweet content matches too many clusters, we add it to array *ManualLabel*. If the tweet does not match with any cluster (meaning it is most likely noise in the data), we add it to an array *Unmatched* to be manually labeled. At the end of the process, we calculate the volume of tweets stored in *EachPoint* array that have been successfully classified under each cluster.

Algorithm 1: Classifying tweets based on cluster

```
Matrix_Cluster MC = {MC_1, ..., MC_N};
Content T = {T_1, ..., T_M};
initialization: Unmatched = array ;
initialization: ManualLabel = array ;
initialization: EachPoint = array[N] ;
for g := 1 to M do  Evolutionary loop
    initialization: Point = array[0] * N ;
    for word in T_g do  Evolutionary loop
        for p := 1 to N do  Evolutionary loop
            if word in MC_p then
                Point += 1;
            end
        end
    end
    maxP = max(Point) ;
    if Point.count(maxP) > 1 then
        append T_g to ManualLabel ;
    end
    else
        if sum(Point) == 0 then
            append T_g to Unmatched ;
        end
        else
            maxIndex = Point.index(maxP) ;
            EachPoint[maxIndex] += 1 ;
        end
    end
end
```

IV. Evaluation and Results

After data collection, we performed our main analysis on tweet volume distribution, content clustering and geographical distribution. We also performed analysis on user account 'mention' frequency and on URL frequency. Below is a detailed summary of our findings.

A. Overview of Tweets

Based on the data we collected, we first compared the total amounts of tweets by date to get initial concept of tweets related to malaria. Our results are seen in Fig. 1. The first search period of tweets was from before our preliminary findings (April 21 to May 3). There is a large spike in tweet volume on April 25 which can be explained by the fact that April 25 is known as World Malaria Day [12]. Between April 25-28, approximately one third of tweets in our dataset were related to World Malaria Day. Many key phrases were more frequent in tweets related to World Malaria Day such as *end malaria* and *malaria day*. When performing other analyses, we would want to consider a normal volume of tweets to have reduced noise. Including such a special event in our evaluation data set could lead us to an invalid conclusion. Hence, we decide to filter out the tweets we collected between April 25 to April 28, and combine the rest tweets from April 21 to May 3 as the "special week" to compare with two other weeks from

May 12 to May 25.

There is another peak on May 23th, yet the peaks volume is not as noisy as the April 25th spike and thus could be an indication of increased spreading of malaria. For our remaining analysis, we focused mainly on the second time period (May 12-25) as well as previous data excluding World Malaria Day interference.

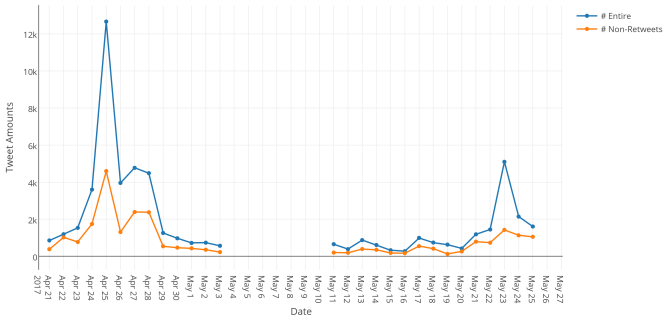

Fig. 1. Volume of Tweets in dataset by day

B. Content Clustering

Using the method described in our *Experiment Setup* section, we use parameters of 20 for total number of clusters for our entire dataset and 8 words in each cluster. We also manually remove some of the irrelevant (noise) words in the cluster. Table I below describes our results.

TABLE I
CONTENT CLUSTER ON ENTIRE DATASET OF TWEETS

Cluster	Results
1	malaria,important,worldmalariaday,higher,disease, geography,tropical,disease
2	air,pollution,kills,globally,children,indoor,yes
3	styled,living,god,self,died,2015
4	free,rise,maldives,murder,travelers,make
5	vaccine,ghana,malawi,2018,sea,created,enemy,trust
6	korea,north,contact,oks,civilian,usatoday,malaria, washington
7	treat,phrase,lots,heard,better,cure,join,prevention
8	oneactagainstmalaria,did,nigeria,know,source,malaria, stopped,estimates
9	bush,george,got,say,field,fighting,film
10	malaria,anti,world,africa,health,day
11	tickets,hotels,combined,more,travel,aids,plane,who,money, malaria
12	cases,plant,therapy,cured,drug,africa,million,malaria
13	scientists,fight,finally,discovered,malaria,vaccine,report, progress
14	drugs,experiences,worst,limpopo,akwa,ibom,provides
15	end,oneactagainstmalaria,menace,africa,ending,malaria, play,efforts
16	courses,mooc,harvard,university,online,free
17	prevention,5000,persons,quicker,route,kick,better,malaria
18	aids,spent,tickets,plane,hotels,health,world
19	experimental,monkeys,protects,deadly,modified,vaccine, malaria
20	fund,budget,million,malaria,trump,cuts,aids

We can see that the clustering method is quite useful since each of our clusters generally represent a different topic regarding malaria. *Cluster 11*, for example, contains a malaria-related news "WHO spent more on plane tickets and hotels than AIDS and malaria last year". *Cluster 19* contains tweets referring to malaria research as "Modified experimental vaccine protects monkeys from deadly malaria". Some of the clusters contain some unrelated noise, such as tweets "Indoor air pollution kills more children than HIV/AIDS, malaria, TB" being assigned to *Cluster 2*, etc. Using this table, we can manually check the most frequent tweets about malaria and cluster each of our clusters into three main categories: News, Research/Epidemic, and Unrelated Noise. For these 20 clusters, our results are shown in the Table II below.

TABLE II
CATEGORIES OF CLUSTERS

Cluster	Results
News	*Cluster 1, 6, 11, 15, 16, 18, 20*
Research/Epidemic	*Cluster 5, 7, 8, 10, 12, 13, 14, 17, 19*
Unrelated Noise	*Cluster 2, 3, 4, 9*

By utilizing this clustering method, we get an initial idea of the contents of our collected tweets. Since we are planning to tracking any epidemic of malaria, we also need to count volume of tweets in each category based on date to see if we can relate this to our observed peak during Week 3 in Fig. 2. Thus, we use *k-means* to cluster our Week 3 tweets.

For each day, we use *k-means* to first cluster the tweets and remove irrelevant words in each cluster. We utilize the Algorithm 1 described in our *Experiment Setup* section to count for tweet volume.

Categorizing tweets into one of our three categories involves a lot of manual labeling. This algorithm relieves us of some of the burden and can quickly count the tweet amounts of each cluster. Running the algorithm yielded the following results seen in Fig. 2.

We see that there are only a small amount of tweets each day that relate to malaria research or spread. This result surprised us and led us to conclude that only tracking volume of tweets will not give us our desired results.

C. Geographical Analysis

Only looking at tweet volume by date does not provide an appropriate evaluation of epidemic. We observed in our clustering analysis that only a small fraction of all tweets were related specifically to an epidemic. Thus, for performing geographical analysis, we decided to look at two specific key phrases: *malaria vaccine/vaccination* and *malaria outbreak* and generate geographical maps.

Our data set is still split on three separate time periods (April 21-May 3, May 12-May 18, May 19-25) in order to

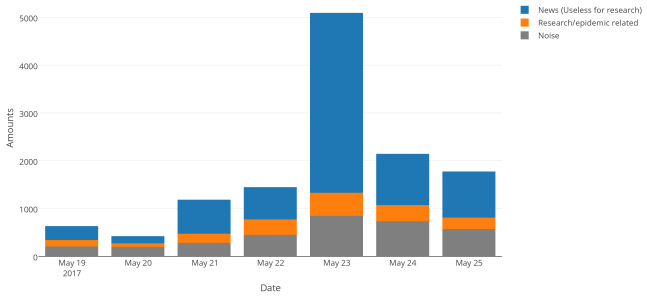

Fig. 2. Content cluster from May. 19 - May. 25

compare patterns between weeks. We first classify Tweets by grouping them as being related to either *vaccination* or *outbreak*. Examining tweets, we found a pattern that for tweets about malaria *vaccine/vaccination*, relevant information regarding the tweets locations usually occurs in label text or location. For tweets about malaria *outbreak*, information comes from in a tweet's content was found to be more useful. This is because a tweet could be sent from a different location but the actual content of the tweet discusses the location of an epidemic (ex: someone from California tweets an article about a malaria outbreak in Kenya). Thus for *outbreak* we use the *geoText* and *geograpy* python package to check if the contents have location information for us to use. We then query those locations to *nominatim* to get a json result and then extract the latitude and longitude for further use.

After all coordinates of the locations were extracted for the *vaccine/vaccination*, we combined the three time periods of tweets together and plot them using *geoplotlib*. Our results are shown in Fig. 3. From here we can see that the tweets are scattered in more developed countries throughout North America and Western Europe as well as in developing countries throughout Africa and south Asia. Our result matches our initial thoughts in that the main research of malaria treatments and vaccines comes from the United States and Europe, while the places where the vaccines are being experimented include African and south Asian regions.

Using *plotly*, we generated a geographical map of tweet volume related to the keyword *outbreak*. Our results are shown in Fig. 4, broken down by time period.

This further confirms to our initial thoughts about the current situation of the malaria epidemic: Africa, South America and South Asia are the major suffering areas. Africa is the most critical region, as we can see from three weeks that the malaria is spreading throughout different countries.

Taking a closer look at Africa, we compare the *vaccine* keyword plot with the *outbreak* keyword plot. This is shown in Fig. 5.

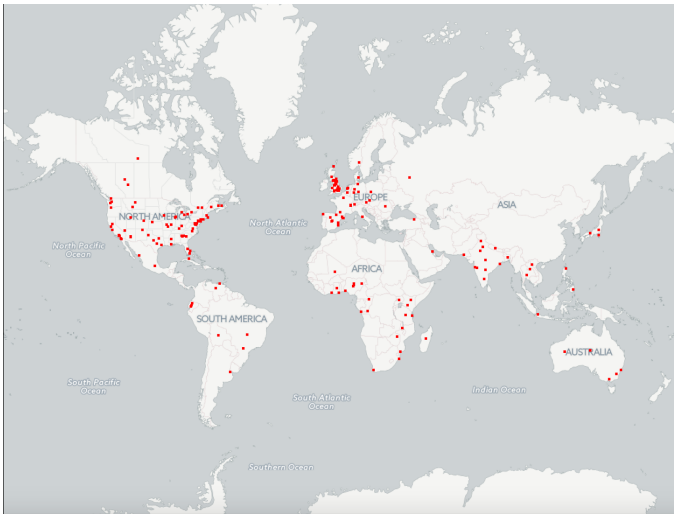

Fig. 3. Labeled locations of Tweets related to "vaccine" and/or "vaccination"

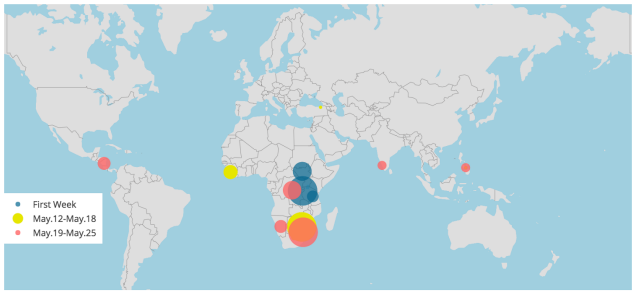

Fig. 4. Labeled locations of Tweets containing "outbreak"

Fig. 5. Zoom in on Africa of Figures 3 and 4

Majority of locations of tweets containing *vaccine* have almost no overlaps with locations of tweets containing *outbreak*, with one exception being Tanzania (the one green dot seen in Fig. 5). Other countries suffering from malaria like Burundi, South Africa, Sierra, Congo, Namibia, and South Sudan have zero *vaccine* related tweets. This could mean that currently vaccines are not being provided to the countries that need them the most to help mitigate the spread of malaria. Such a result is

worth investigating further since this means that the countries most suffering from malaria should get more attention from organizations such as the World Health Organization. For future work, including a keyword list with other malaria-prevalent countries for data collection could assist in tracking of malaria spreading situations.

D. Other Overall Analysis

1) Mentioned Accounts Analysis: Beside the analysis explained above, we also investigated the frequency of specific Twitter accounts mentioned over the course of the three weeks. This would reveal to us which organizations are at the forefront of attention for malaria on the Twitter platform.

As we can see in Fig. 6, the top 20 mentioned accounts happen to be mentioned over 100 times, many of them being retweets. We analyzed distinct contents of the top 20 accounts and then clustered them according to the topic of their contents. Our results are shown in Fig. 7. Interestingly, only 24.4% of those top tweets were related to malaria research or epidemic. 32.7% were related to news about malaria and 42.9% are noise. Looking within the tweets containing news contents, there were three main topics being discussed. The majority of tweets pertained to Donald Trump's budget cuts, the World Health Organization spending too much money on travel as well as the WHO leader election. News and spread of awareness of malaria is useful and helps to explain the two major volume peaks we observed earlier in Fig. 1. Nevertheless, these tweets do not reveal enough information involving research or the spread of malaria. Hence, we can conclude that if someone who intends to analyze malaria for research purposes and to monitor epidemic situations, following these top 20 accounts wont help much.

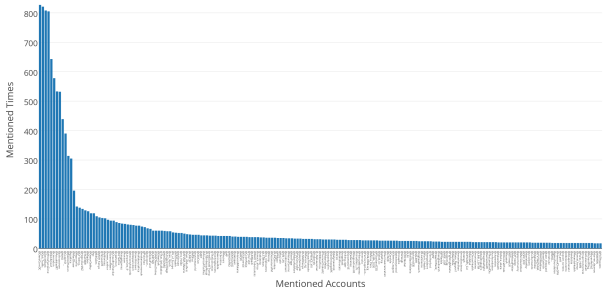

Fig. 6. Top 200 mentioned accounts related to malaria

2) URL Frequency Analysis: Of the tweets analyzed in our dataset, 2276 contained URLs (around 8% of all collected tweets) with 1502 being unique shortened URLs. However, it must be noted that mant of these were shortened and actually linked to the same content or web page.

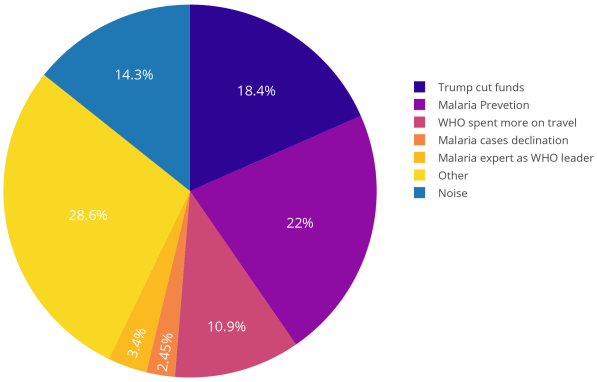

Fig. 7. Clusters of tweet categories of tweets from Top 20 mentioned accounts

Of the tweets containing URLs, there were 19 URLs that were contained in 10 or more separate tweets. After processing and eliminating duplicate tweets and tweets that provided URLs that did not link to an actual webpage, we were left with 9 URLs that were mentioned in 10 or more separate tweets. Looking at those URLs revealed three main topics related to malaria that were being tweeted about: 1) the World Health Organization (WHO) spending more money on travel expenses than on fighting malaria, 2) malaria prevention and treatments and 3) Donald Trumps budget proposal reducing funds to fight malaria. Both Table III and Fig. 8 show our results broken down into each of these categories. The most tweeted URL by far was an article published on May 23 from Quartz entitled World Health Organization spent more on plane tickets and hotels than AIDs and malaria, mentioned in 200 tweets [11]. Of all the tweets containing URLs, 10.59% of them linked to articles talking about the WHOs spending.

TABLE III
TOPICAL ANALYSIS OF URLS MENTIONED IN $\geq$ 10 TWEETS

TOPIC	WHO Spending	Malaria Prevention/ Treatment	Trump Budget Cuts
Quartz Article	200		
Eldiario Article	16		
UNWatch Tweet	14		
NBCNews Article	11		
ScinceNews Article		25	
A. Gabaldon Article		11	
PBS Vaccine Article		10	
Reuters Tweet			18
NickKristof Tweet			12
Total Tweets by Topic	241	46	30
Percent of All Tweets with URLs	10.59%	2.02%	1.32%

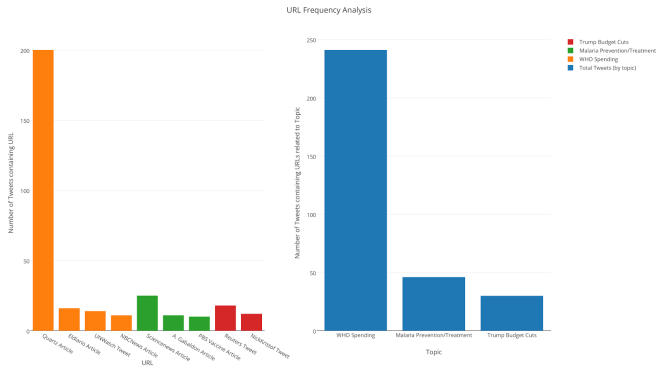

Fig. 8. URL Frequency Analysis by Topic

V. Conclusion and Future Work

Tracking the malaria epidemic seems to be possible using our final defined list of key phrases. Through analysis, tracking the amount of tweets by day did not provide enough information to understand the epidemic situation as too many tweets pertained to news as opposed to research or outbreak. After applying the *k-Means* method with our algorithm, its clear that for each day, there is only a small amount of tweets related to research or epidemic. Creating a geographical map on our entire dataset did not reveal very useful results. However, we discovered from tracking the specific key phrases *vaccine/vaccination* and *outbreak* that there is almost no overlap between the locations of the tweets that contain these two keywords. This led us to conclude that the areas suffering worst from the disease may not be getting enough assistance as it seems vaccines are not being provided there.

From our analysis of mentioned accounts and URL frequencies, we conclude that the most frequently mentioned accounts did not give information on the research or epidemic domain of malaria. Many of the tweets from these popular accounts referenced news, confirmed from our content clustering results.

For future work, we plan to improve our algorithm for content clustering. We desire a better method for labeling tweets as manually labeling tweets is very inefficient. Following the geographical analysis, researchers should look into monitoring tweets from countries with high malaria prevalence in order to better determine a relationship between social media and the spread of the disease. Such research will require a long period of data collection. Yet we think it will help a lot in tracking the spread of malaria, hopefully one day leading to its elimination.

References

[1] World Health Organization. ”Fact sheet about Malaria.” World Health Organization. 24 Apr. 2017. Web. 5 Jun. 2017. <http://www.who.int/mediacentre/factsheets/fs094/en/>

[2] BBC News. ”Malaria: Kenya, Ghana and Malawi get first vaccine - BBC News.” BBC News. n.d. Web. 5 Jun. 2017. <http://www.bbc.com/news/health-39666132>

[3] Malaria No More. ”Funding the Malaria Fight — Malaria No More.” Malaria No More. n.d. Web. 5 Jun. 2017. <https://www.malarianomore.org/advocacy/funding>

[4] McNab, Christine. ”What social media offers to health professionals and citizens.” Bulletin of the world health organization 87.8 (2009): 566-566.

[5] Middleton, Stuart E., Lee Middleton, and Stefano Modafferi. ”Real-time crisis mapping of natural disasters using social media.” IEEE Intelligent Systems 29.2 (2014): 9-17.

[6] Signorini, Alessio, Alberto Maria Segre, and Philip M. Polgreen. ”The use of Twitter to track levels of disease activity and public concern in the US during the influenza A H1N1 pandemic.” PloS one 6.5 (2011): e19467.

[7] Missier, Paolo, et al. ”Tracking Dengue Epidemics using Twitter Content Classification and Topic Modelling.” International Conference on Web Engineering. Springer International Publishing, 2016.

[8] Lee, Kathy, Ankit Agrawal, and Alok Choudhary. ”Real-time disease surveillance using twitter data: demonstration on flu and cancer.” Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2013.

[9] Pedregosa, Fabian, et al. ”Scikit-learn: Machine learning in Python.” Journal of Machine Learning Research 12.Oct (2011): 2825-2830.

[10] Kriegel, Hans-Peter, Erich Schubert, and Arthur Zimek. ”The (black) art of runtime evaluation: Are we comparing algorithms or implementations?.” Knowledge and Information Systems (2016): 1-38.

[11] Chutel, Lynsey. ”World Health Organization spent more on plane tickets and hotels than AIDS and malaria.” Quartz. n.d. Web. 5 Jun. 2017. <https://qz.com/989819/who-spent-more-money-on-travel-than-aids-malaria-ap-investigation-reveals/>

[12] World Health Organization. ”World Malaria Day, 25 April 2017.” World Health Organization. 24 Apr. 2017. Web. 5 Jun. 2017. <http://www.who.int/campaigns/malaria-day/2017/en/>