Power Outage Prediction via Twitter

Rodolfo Santana

Computer Science Capstone Project 499

Advisors: Chris Fernandes, Nick Webb

March 17, 2016

Abstract

There has been numerous amount of research done where researchers have mined social media in order to predict a variety of economic, social, and health related phenomena[2]. What I hope to accomplish is fine-grain prediction of outages across the US by monitoring tweets that either report or hint of one occurring. Data mining Twitter for information and using machine leaning methods to attain valuable data, I was able to evaluate roughly 5,000 tweets collected from December, 2015 though February, 2016. The system created showed that I am able to detect actual outages occurring at their precise location. There is still work to be done to be able to predict outages in a much more granular level, but as of now, some level of prediction was attained.The system plots cascading-failures as they unfold in real-time, which helps as a predictor of outages.

Contents

List of Figures

1 Introduction

It is no surprise as to how much social networks have become part of our everyday lives. Everyday millions of people log on to these networks and generate tremendous amounts of data. Since its launch in June of 2006[7], Twitter has become one of the most popular social networks today. With over 500 million users, 302 million of those being active, Twitter generates about 340 million tweets a day[7]. Due to Twitters rapid growth in popularity, researchers have used this social network in order to try and predict things such as the stock market, or the spread of disease[2].

One never-ending problem that people from all parts the world face are power outages. A power outage is defined as a short- or long-term loss of electric power to an area. There are numerous causes for a power outage just like there are a number of effects caused by them. What many do not know is the threat power outages have to critical infrastructure such as telecommunication networks, financial services, hospitals, etc is tremendous. From a business perspective, power outages are one of few business disruptions that companies, more often than not, are unprepared for[6]. That being said, they are unaware of the true costs and impact that they can have on their operations[6]. From a national security perspective, power outages heighten the risk of terrorist or cyber attacks[6]. The effects a power outage can have are endless and in many cases may not be as dire as the examples discussed. Nonetheless, the occurrence of power outages are a problem that effect people worldwide.

As a result of this global problem, the following report discusses the development of a system that combines both data mining Twitter and machine learning that extracts latent information from tweets in order to answer the question: Can we use Twitter as a predictor of where power outages will occur?

2 Background and Related Work

Ample amounts of work have been done in the past in where Twitter was utilized as a predictor. Most recently, Dr. Adam Sadilek, professor at the University of Rochester, developed a system in where Twitter was used in order to predict and model disease transmission and spread. Dr. Sadilek began his work by collecting tremendous amounts of data from Twitter. Utilizing a Search script, he periodically queried

Twitter for the most recent tweets talking about the flu or flu-like symptoms within 100 kilometers of the New York City city center[1]. Once the data had been collected, he developed a language model that would classify individuals as either "not sick" or "sick" based on the text of the tweet[1]. From there he was able to map out where in New York City, sick people were located and compute the likelihood of someone getting ill by monitoring where they went and who they came in contact with.

Ultimately, Dr. Sadilek was successful in developing a system that shows the spread of disease and computes the likelihood of someone getting sick. His work served as a guide as to how to approach the problem which this paper explores. Similar methods used by Dr. Sadilek's were also used in this project in order to try and create a system that would successfully predict power outages.

3 Why Twitter?

Before going into further detail about the project you may be wondering why Twitter? Why not mine other social networks such as Facebook? Aside from its obvious popularity and the amount of data it generates, the way in which user messages are structured on Twitter is a unique feature. User messages, or tweets, have a limit of at most 140 characters. As a result, it is easier to look for data that will best help answer the question: Can we use Twitter as a predictor of where power outages will occur? This forced brevity forces the user to think carefully about the way in which they word their message. Carefulness of word choice is a vital part of the project. As you will see later on in the report, the brevity of tweets will make it much simpler to look for information-rich and relevant data.

4 Motivations

Over the winter break, I had the opportunity to spend some time with crew members of Southern California Edison's repair team. Southern California Edison is the primary electricity supply company for much of Southern California. From my conversations with them regarding outages, I was found that detection and restoration time were two fields which they always seek to improve. In some cases they are unaware that an outage has occurred either because people have not reported it to Edison, or their phone lines are backed up so it takes them a while to become aware. These problems were in part my motivation for developing a system to solve this problem. The repair team were intrigued with the idea of using Twitter as an avenue for improving reaction and restoration time. The idea of predicting outages before they even happened was an idea that was a bit hard to grasp, but nonetheless fascinating.

5 Project Overview

Figure 1 is a bird's eye view of the project created in order to answer our research question. The first step in the project is collecting data. We do so using a Search script in order to search for data within Twitter. Before storing the data in a database, I filter out the irrelevant data collected and classify each piece of data into one of two groups, ~~groups~~ discussed in further detain in Section 7. Once stored in the database, I am able to plot the data with the visualizer. The visualizer is a ~~visual~~ tool that will ultimately help find patterns within the plotted data. These patterns will then help towards prediction, which is the ultimate goal. The following report will explore each part of the project, what it does, how it works, and its importance.

6 Data Collection

As mentioned earlier, the amount of data generated from Twitter is enormous. In mining this mountain of data, one can easily get stuck with useless information. Therefore we use the Twitter Search API in order to refine the search of tweets. The Search API allows queries against the indices of recent or popular Tweets[4]. It behaves similar to the Search feature available in Twitter mobile or web clients, but allows the user to refine the search in much more detail. For the purpose of this project there are really two things we care about, is the data relevant and where is the user when tweeting. The following section will discuss how I attained relevant data and data that has the user's location.

6.1 Is the Data Relevant?

When talking/reading about power outages, certain words and/or phrases come up quite often, for example, brownout, blackout, power outage, etc. I first began by searching these terms utilizing the Search feature on

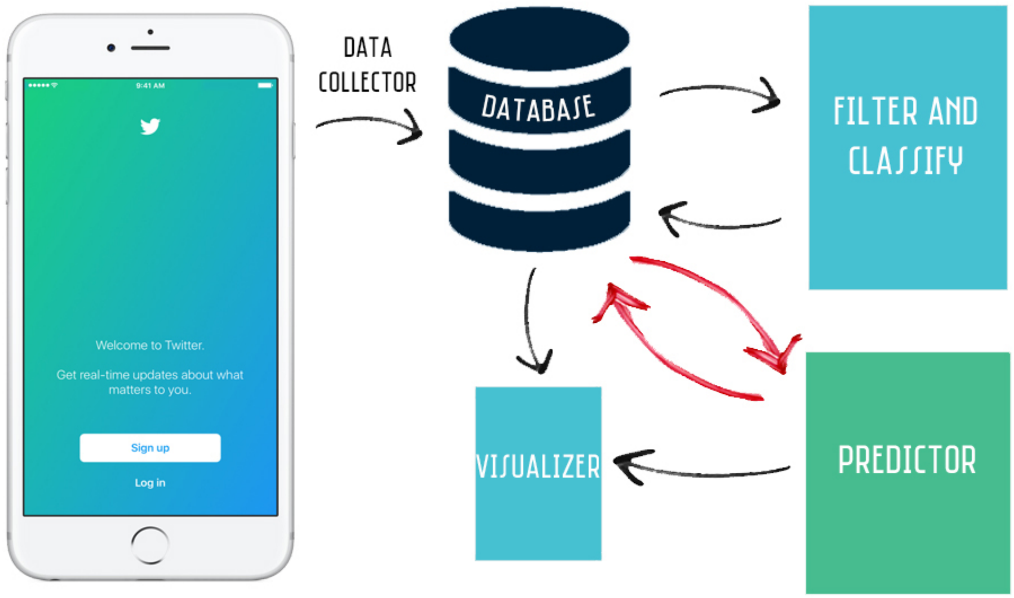

Figure 1: An overall view of the system. ~~The system comprises of~~ four parts: data collector, database, filter and classifier, and a visualizer.

the Twitter web client. This gave me a general idea of the types of tweets associated with these words. For example, brownout is an important term when talking about power, but surprisingly, people do not tweet brownout as often, and if they do, they may not be referring to the type of brownout relevant to power.

After doing some research and cross-checking terms with the Search feature on the Twitter web client, I generated a list of terms that would provide tweets relevant to power. This list of terms was refined once more after carrying out a survey. This survey asked people to write down terms they would use if they were to tweet about a power outage. After analyzing the two lists generated by myself and the survey takers, the finalized list of terms to look for was the following: *power outage*, *power failure*, *power blackout*, *powers out*, *no electricity*, and *flickering lights*.

6.1.1 Filtering Out Irrelevant Data

Although the terms looked for provided tweets containing the terms, that did not necessarily mean they were relevant to the topic. When analyzing tweets, I found that the majority of them that separated the

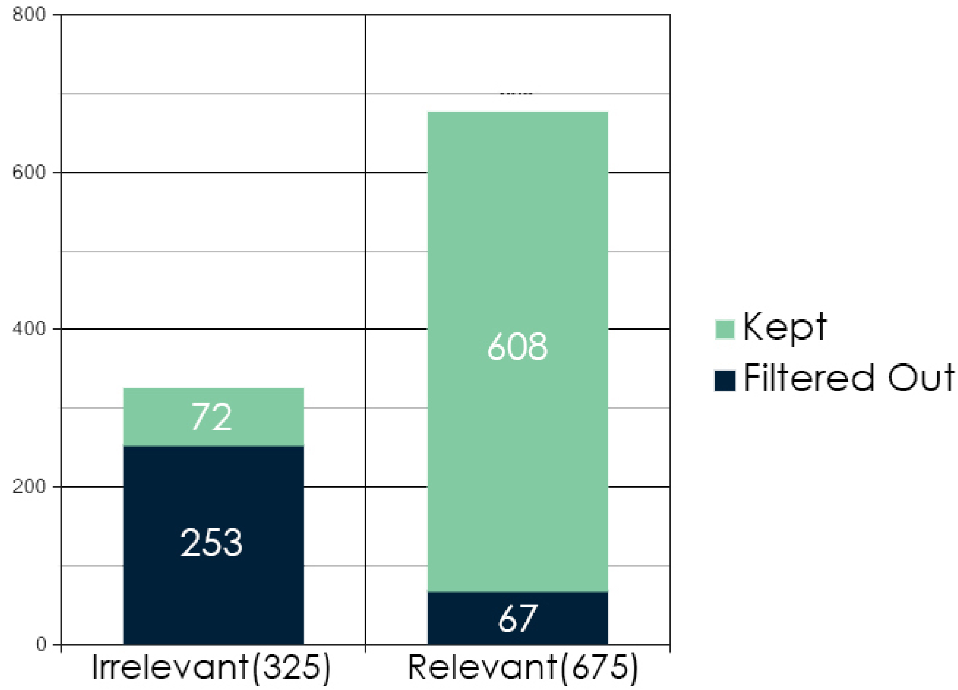

Figure 2: Bar graph represents the effects of the filter primarily used to filter out irrelevant tweets.

term were irrelevant. By that I mean that the term *power outage* would be in the tweet, but there would be other words in between *power* and *outage*. The majority of the tweets collected that had the words separated throughout the text were filtered out. The only exception in where I did not care about the separation of the term within the text was for the term *flickering lights*. For this case, filtering out tweets that separated *flickering* and *lights* filtered out more relevant than irrelevant tweets.

Figure 2 shows the results the filter had on a set of 1000 collected tweets. From the bar graph, we can see that the filter filtered out both relevant and irrelevant tweets. Despite the fact that we lose roughly 9% of relevant tweets and keep 22% of irrelevant ones, the most important thing to note is that 78% of irrelevant tweets were successfully filtered out. Despite some false positives and false negatives, we can be sure that the majority of the data being collected by our Search script is talking about outages.

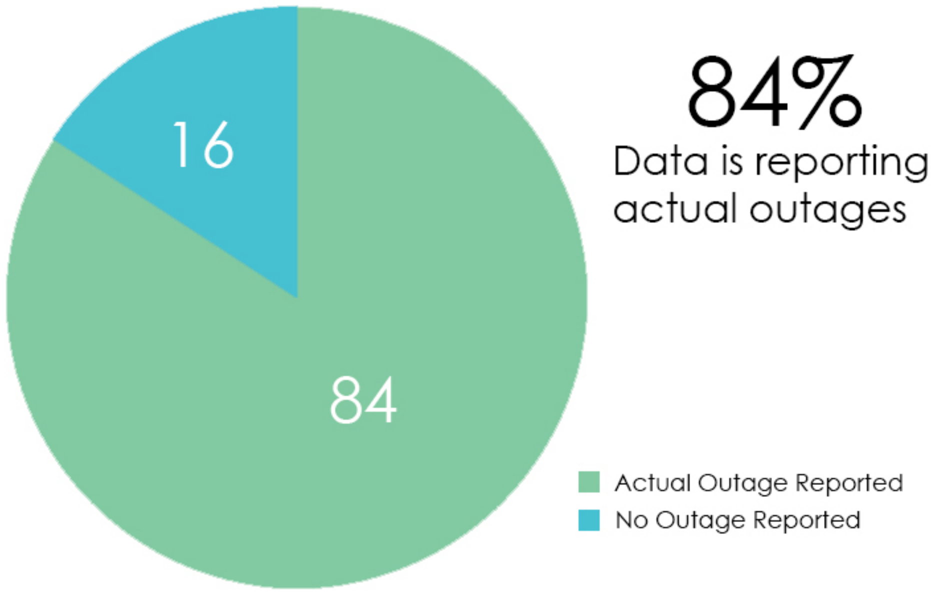

Figure 3: Pie chart represents tweets that reported actual outages vs tweets that claimed there was one, but there was no confirmation of it occurring.

6.1.2 Is the Data Reliable?

Once tweets were retrieved from Twitter, the question that was left to answer was determining whether or not these tweets were true. Just because someone tweeted about a loss of power, did that mean there was really an outage? Figure 3 shows the results of a test carried out in where one-hundred of the searched tweets were analyzed. In this test, I checked the outage reports of companies that handled the power supply for the tweets location to see whether or not the data collected was a reliable source of outage detection. Results showed that 84% of the tweets looked at reported actual outages. The other 16% were instances where the tweets reported an outage, but reported one at a different location or companies were not aware of an outage in that location yet. With these results, I was confident that the data collected would help formulate an answer to our research question.

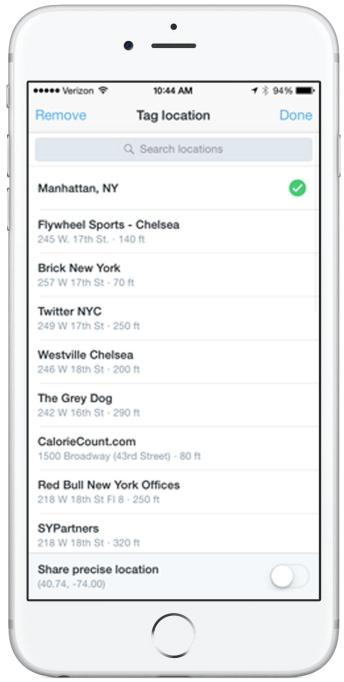

Figure 4: Screenshot a user tagging their location on their tweet. Twitter provides a list of locations to choose from.

6.2 What is the Users Location?

Along with relevance, one crucial piece of information is knowing where these tweets are coming from. Therefore it is important for the tweets being searched to not only be relevant, but be geo-tagged as well. For a tweet to be geo-tagged means that the user has tagged their location from which their tweeting from.

6.2.1 Getting the Precise Location

Just because a tweet is geo-tagged, does not necessarily mean we can get the users precise location. As you can see from Figure 4, when the user chooses to tag their location, they are provided with a list of popular tags within his/her surrounding. Once the user has selected a location, Twitter classifies the tagged location under one of four categories: *Admin*(State), *City*, *Neighborhood*(Gated Community), or *POI*(Point of Interest). POI's are specific locations such as the Bank of America on State Street or a restaurant.

The category determines the level of granularity of the location, with Admin being the coarsest and POI

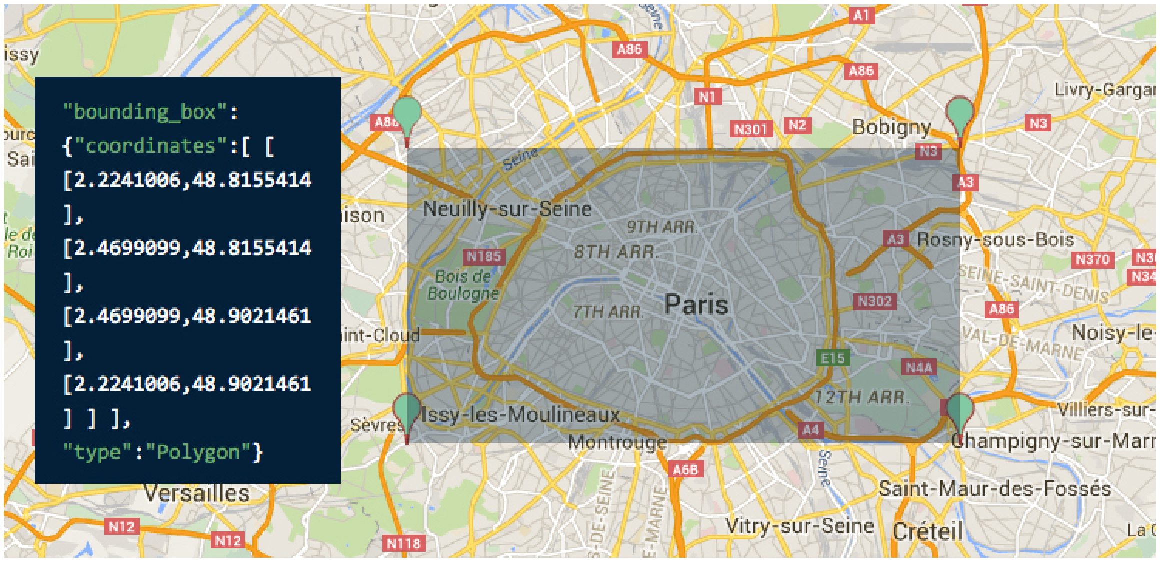

Figure 5: The box on the left represents the bounding box given by Twitter when the user decides to tag Paris as the location. The map represents the area covered by the bounding box.

being the finest grain. Whatever category the tagged location may be, Twitter provides a bounding box composed of a set of four longitudes and latitudes. Figure 5 is an example of a bounding box created when a user tags their location as the city of Paris. As you can see, the longitudes and latitudes are not the same. What these coordinates do is allow for us to see the area in which the user resides according to the tag. Therefore it can be as vague as the city of Paris or as specific as the stadium within Paris, the Parc des Princes. That level of specificity is just what is needed for this project. Therefore, the category of geo-tagged tweets that I am interested in are the tags that fall under the POI category, Point of Interest. Since POI's are a precise location, the bounding box generated contains a set of four longitudes and latitudes that are the same.

Another option the user has when tagging their location can be seen at the very end of the list from Figure 4. The user has the ability of sharing their precise location. Regardless of the category in which the tag falls, by choosing to share their precise location, we are given a bounding box with a set of identical longitudes and latitudes. Therefore, the geo-tagged tweets in which we are interested in are the tags that fall under the POI category or the case in which the user has chosen to share their precise location.

6.3 Where Does the Data Go?

CouchDB is an open source database that focuses on ease of use and on being a database that completely embraces the web. It is a document-oriented NoSQL database that uses JSON to store data[5]. One of the nice features of couchDB is that the data is presented in a very friendly way. CouchDB is very helpful in organizing such data. Therefore, once a database is created, we simply utilize its update function in order to store tweets being collected. CouchDB makes it very accessible to pick out different entities within a tweet which is why I chose to store the data in couchDB.

At this point of the process, the data being collected and stored in the database satisfies the two things I was initially concerned of. First, is the data relevant and reliable, and second, do we know the user's precise location.

7 Data Classification

After collecting tweets, I noticed that a number of them were not reporting outages, but either complaining of light quality and/or shared concern of their power potentially going out. This led to the creation of two categories, PL(power low) and PO(power out). A PL, power low, means that there is not an outage yet, but there is one bound to happen. PL's are also complaints about light quality. A PO, power out, is a tweet in which an actual outage is reported. Each tweet collected would fall under one of these categories, but this led to the following question: How was I going to classify the tweets?

7.1 WEKA

WEKA is a popular suite of machine learning software that is used for predictive modeling. By providing WEKA with an arf file composed of tweets and their classification, classification that was determined to what I understood to be an instance of a PL and a PO, WEKA is able to develop a language model that could classify a tweet as one of the two. After testing a number of data analysis algorithms, the algorithm with the highest percentage of correctly classified instances was a decision tree based algorithm. The algorithm had an effectiveness of 89.2353%.

Figure 6: Example of a tweet categorized as a PL, power low.

The way in which the decision tree algorithm works is by first building a tree from a set of classified samples. Each node within the decision tree is an attribute of data, an attribute that most effectively splits the set of samples into subsets enriched in one class. Once it has reached a leaf node, that is the classification in which the tweet is labeled as.

Figure 6 and 7 are both examples of tweets from our data. Figure 6 is an example of a tweet categorized as a PL, power low. In this case the user shows concern of a potential power outage occurring due to weather conditions and light quality. In Figure 7 it is clear that the user has suffered a loss of power so this is an instance of a tweet classified as a PO.

8 Visualization

Although it seemed great to have relevant data that either reported an outage or warned us of one happening at a precise location, this data is useless unless I were to have one form of representing the data visually. Therefore, using the Google Maps API I was able to plot out the PLs and POs in their given precise location. This form of plotting out tweets would prove be a useful tool in what is the ultimate goal of prediction.

Figure 8 shows the product of the visualizer in the system. The pins are each individual tweets created at that precise location. As represented by the key, PL's and PO's are distinguished by the difference in

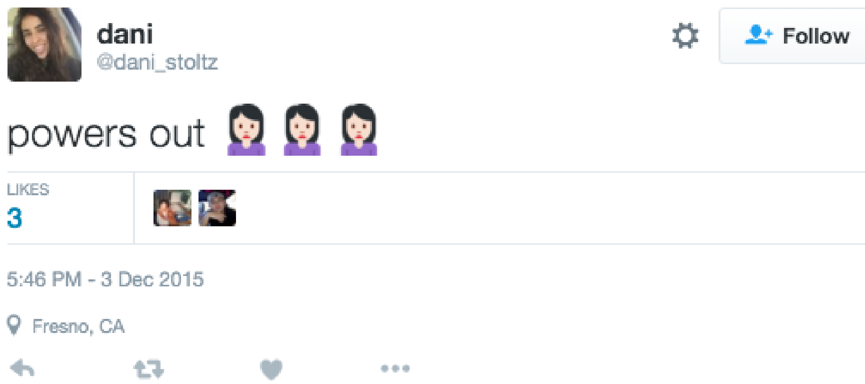

Figure 7: Example of a tweet categorized as a PO, power out.

color. Another feature added to the visualizer is the heat-map effect that is depicted by the gradient of color underneath the pins. The heat-map depicts the density of tweets originating from different locations. In this case, red indicates a higher number of tweets being tweeted from that location.There are three options a user can choose to plot tweets through the visualizer: Day-to-Day, Real-Time, and Group.

8.1 Day-to-Day

The Day-to-Day option is pretty self explanatory. Once picked, the user must select a start and end date. The system then looks in the database for tweets that were created between that time frame. Once it has found all the tweets that fall in the start and end dates given, and everything in between, the system plots out all the tweets on a map similar to Figure 8.

8.2 Real-Time

The Real-Time option is a bit more interesting. This implementation plots out tweets similar to the Day-to-Day option, but instead the system queries the OS for the present day. Once passed in, the stream is initiated. The stream connects to twitter and grabs any tweets that fit the specifications discussed in Section 6. After every ten minutes, the map is updated, plotting any new tweets. This option runs for as

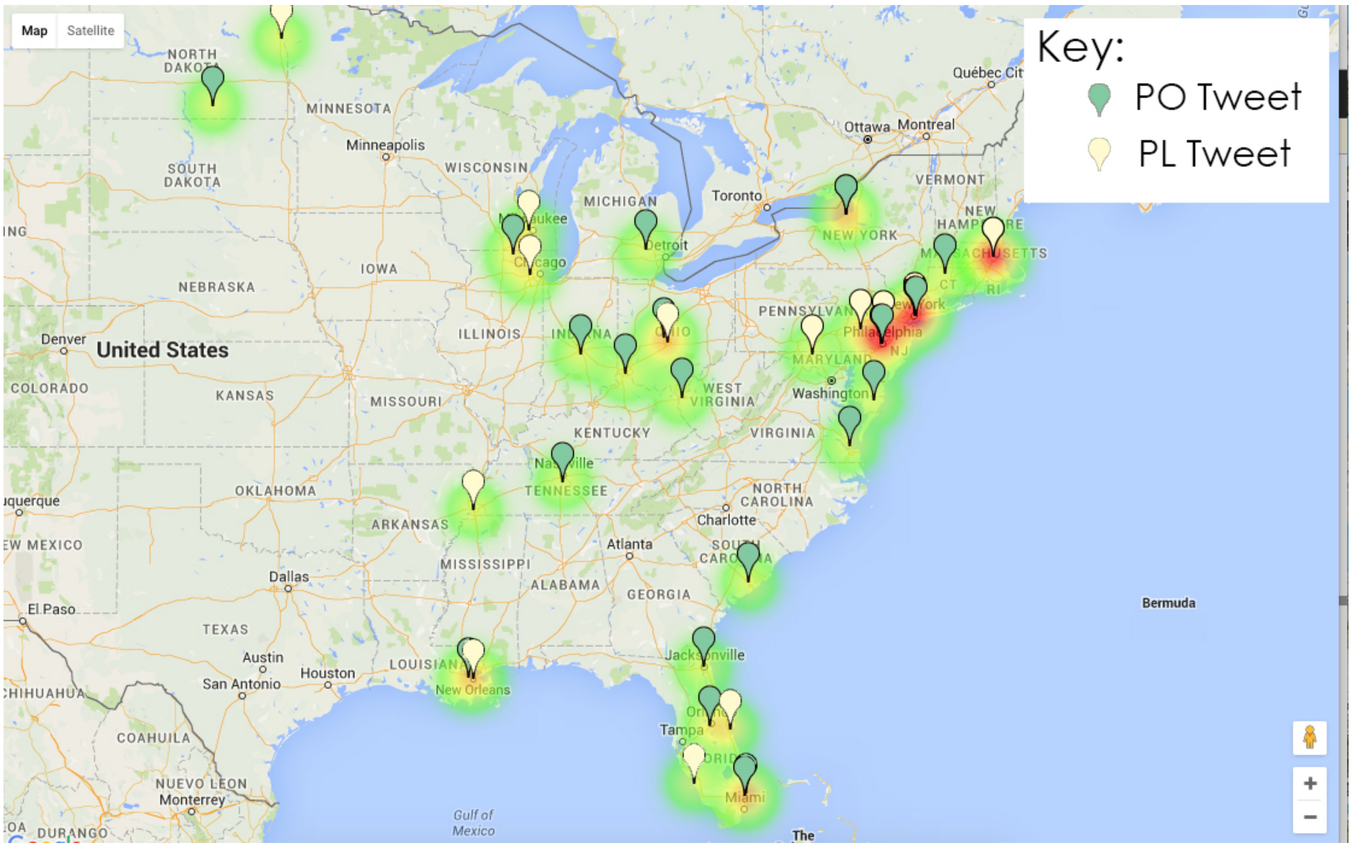

Figure 8: Tweets about power outages across the US from Jan 01 to Jan 10.

Figure 9: Screenshot of a locale. Following tweets were within the given limitations, in this case 25 mile distance of one another and within a three hour time frame.

long as the user would like.

8.3 Group

The last option the user has the opportunity to choose is the Group option as shown in Figure 9. What this option does is group tweets that are close in space and time. Once this option is chosen, the system asks the user to for a distance in miles, X, and for a time Y. The system then takes the specifications prompted by the user and generates groups of tweets that are are most X distance (miles) from one another and that are created within a Y time-frame (minutes) of one another.

Again, the visualizers ultimate goal is to serve as a tool for determining prediction of power outages. What I hoped to get out of plotting PL's and PO's is a visual representation of different patterns that these tweets may have. Not until these patterns are discovered can we begin to answer the question of whether we Twitter can act as a predictor of where power outages will occur.

9 Prediction

Using the visualizer as a tool of mapping out the tweets there was one pattern that seemed most effective for predicting where an outage was heading. From running tests on the system, I found that the Group option was the most effective form of predicting. After numerous trial and error tests of distance and time, limitations of 25 miles and 3 hour time-frame seemed to produce the most useful grouping of tweets.

Once I ran the Group option with the stated limitations, a total of 14 locales were produced. Out of the 14, 9 of them showed a similar pattern as the one illustrated in Figure 10. What I found in those 9 groups was that there was a directional pattern in the way people tweeted about warned of a power outage. As one can see in Figure 10 there is a southward direction in the time people tweeted. The is an outage at 2:11 PM and within the span of an hour there are four other tweets reporting a loss in power and one reporting a complaint on their light quality.

From the directional pattern, I then got a sense of which direction an outage is most likely to occur next. By knowing the direction, we are able to predict an area in which a power outage will happen next. The downfall with this form of prediction is that the level of granularity in which I am able to predict is not as specific as I would love it to be. Although the level of specificity is a bit broad, seeing the directional pattern unfold in Real-Time is very interesting to say the least. When testing the system, that is something that caught my attention right away.

The example discussed above showed that prediction is possible to some level of granularity. From this the following algorithm was developed to predict outages:

Predicting Outages Algorithm

1: Run REAL-TIME option
2: Plot current tweets happening at start time

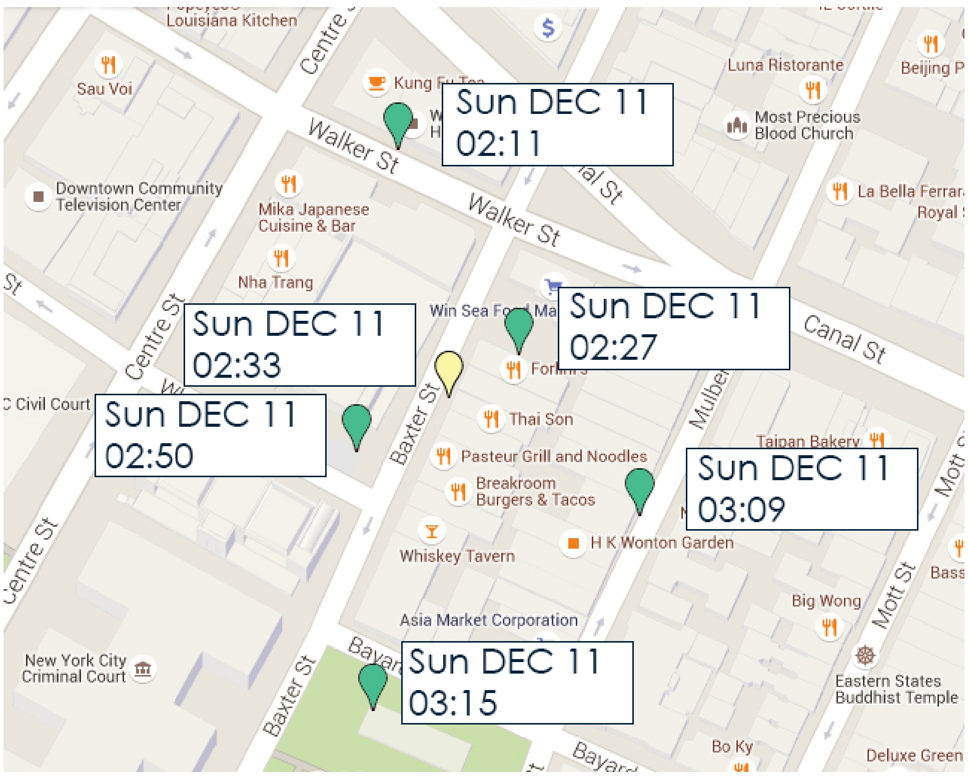

Figure 10: Example of the directional pattern found within groups. Tags next to the tweets represent the time tweeted.

```
Initiate stream
while ESC key not pressed do
    Stream for new tweets
    for Every 10 minutes do
        Run Group method
        Update and plot groups created on map
        if There are groups then
            Notify user "Cascading Failure Detected"
            Notify user the direction and distance outage most likely to occur from groups most recent tweet
        end if
    end for
end while
```

The algorithm describes a Real-Time implementation incorporating the Group method. This will help determine prediction of outages that have yet to occur. The algorithm begins by plotting PL's and PO's that have been tweeted since at that present time. After every ten minutes, we look in our database for group of tweets that are within a 25 mile radius of each other and tweeted within three hour time frame. Once a group is found, we are able to see a direction in which the outages or warnings are occurring. Therefore, the system can predict and warn the user the direction and distance in where the next outage will occur.

10 Conclusion

There are a number of features this system is capable of performing. One of those features is detecting potential and already occurring outages. As talked about in Section 6, 84% of the time a tweet is talking about an actual outage. Therefore I am able to plot out where these outages are occurring all throughout the US through the visualizer. This can be done by both searching tweets in the past or streaming for real-time tweets.

Secondly, another implementation of the project is making clusters of tweets. These clusters are formulated by collecting tweets that are close to one another according to their distance and time. This allows for the user to see the area affected by an outage. This Group implementation lead to the prediction of outages by means of direction. Seeing a directional pattern within a group of tweets, I am then able to predict a general area in which an outage is most likely to occur by following the direction of the group.

We asked early in the entire process whether Twitter could be used as a predictor of where power outages will occur next. After going through the entire process in which this paper illustrates, I am confident that we are able to predict outages via Twitter, but to some extent. The ultimate goal is being able to predict the precise location of where an outage will occur next. This is only one step towards achieving that overall goal in hopes of preventing unnecessary instances of this global problem.

11 Future Work

As stated early on in the paper, motivations for this project sprung from interviewing crew members from the repair team at Southern California Edison. For them detection and reparation time has been an area in which they have always seek to improve on. Seeing that my current system is a strong detector of power outages, I would like to test to see if companies can improve detection and restoration times by utilizing it. Also, I would like to make the visualizer user interactive. When a power outage location is chosen, I would like for automatic tags and labels to pop up, containing useful information about the tweet.

A tweet is composed of a number of entities. As it stands and as discussed throughout the paper, the entities explored within a tweet were: text, location, and time created. Perhaps other entities such as the number of likes/retweets a tweet has can be a better form of prediction. Does the number of likes/retweets give that tweet/user more credibility? Does that imply that other users had an outage as well but rather than tweet it they just liked/retweeted another users tweet? Furthermore, it would be interesting to see what other users are Tweeting about when there is an outage. Perhaps that is something interesting in its own. Ultimately the goal is to be able to give a precise location of where an outage will occur, this is only one step towards achieving that goal.

References

[1] Adam Sadilek, Henry Kautz, Vincent Silenzio, *Predicting Disease Transmission from Geo-Tagged Micro-Blog Data*, Twenty-Sixth AAAI Conference on Artificial Intelligence, 2012.

[2] Adam Sadilek, Henry Kautz, Vincent Silenzio, *Modeling Spread of Disease from Social Interactions*, Sixth AAAI International Conference on Weblogs and Social Media (ICWSM), 2012.

[3] Sean Brennan, Adam Sadilek, Henry Kautz, *Towards Understanding Global Spread of Disease from Everyday Interpersonal Interactions*, Twenty-Third International Conference on Artificial Intelligence (IJCAI), 2013.

[4] The Search API, https://dev.twitter.com/rest/public/search, Accessed: 2015-11-19.

[5] CouchDB, https://en.wikipedia.org/wiki/CouchDB, Accessed: 2015-11-19.

[6] Allianz, http://www.agcs.allianz.com/insights/expert-risk-articles/energy-risks/, Accessed: 2016-2-25.

[7] Twitter, https://en.wikipedia.org/wiki/Twitter, Accessed: 2015-11-19.