Hype Detection in Twitter

Yiming Guo
Computer Science Department
Illinois Institute of Technology
Chicago, Illinois 60616
Email: yguo46@hawk.iit.edu

Aron Culotta
Computer Science Department
Illinois Institute of Technology
Chicago, Illinois 60616
Email: culotta@cs.iit.edu

***Abstract*—This project focuses on the Hype Detection in Twitter. Hype can be considered as promoted tweets, which being paid by a third party then showed up on the users' homepage. On the contrast, genuine tweets are common tweets which generated by normal accounts and involved zero money deals. The main method of distinguishing promoted tweets is RandomForest.**

Keywords*—*Individual Research, Machine Learning, Hype, Twitter, Social Media.

I. Introduction

This project is about Hype Detection on Twitter. Hype means the kind of topics or key words which become popular because someone or some companies made them be. Usually by paying money.

The idea of Hype Detection is inspired by the song Gangnam Style. This song was so famous at 2013 but in a talk show Morning Call, which is considered to be frank and authoritative in China. The host is Xiaosong Gao, a director and writer. He, as a part of Hollywood guys, said this song was a hype made by a company in U.S. He said he knew someone in that company and knew all the details beneath the ground. No matter his words are trustable or not, we all may have noticed once or twice that sometimes a person or an event becomes well known in a tiny period of time. That is the kind of hype we wanted to look for.

However, there's no way to precisely distinguish whether a tweet is a hype or not. For example, many experts thought that Gangnam Style was a hype because the song itself was not good enough to become the best one in every music board (such as Billboard in the US). But what criteria can be told by observing this event and claim that it's a hype? Nothing!

The ultimate goal of this project is detecting such hypes beneath the ground without any clear labels. For now, the starting point is identifying the clearly labeled promoted tweets as an easier goal.

II. Data

The data being used in this project comes from Twitter. In this project, promoted tweets which showing all the time on Twitter user's homepage is the target. Figure 1 gives an example of promoted tweets.

On the other hand, genuine tweets are the normal tweets which could be posted by anyone. In this project, genuine tweets are being selected from all around the area and sources, such as politics, sports, daily lives, from CNN, from ESPN, and from normal guys. Figure 2 shows an example of genuine tweets.

Figure 3 shows a clear label of promoted tweet promoted by Starbucks. Unfortunately, this label is not available in the status of tweet returned by Tweepy (a library of Twitter developer tools). It means, even if this tweet is showing on user's homepage with a label called promoted, it's still unable to catch a status label returned by Tweepy and say "It's a promoted one". Manual method is not enough for scientific research, so it's necessary to find another way out.

Fig. 1. An example of promoted tweets

Fig. 2. An example of genuine tweets

Fig. 3. An example of clear labels of promoted tweets

III. Method

The method finally being used is Random Forest classification. Since there's no easy way out, we need to build up a training model and let it tell the difference between two kind of tweets.

There're 17 parameters in this classification: minimum, maximum, and mean for Jaccard similarities of users' friends between each user pair, 10th quantile and 90th quantile. The user pair means pairs of users between each two retweeters in a single tweet. Because therere 3 kind of pairs, including friends, followers, and a set of combination of friends and followers. So the number of features must time 3. So its 15 now.

For example, there're around 300 tweets labeled as promoted tweets in my database. For each tweet, there may be 100 retweeters who retweeted this tweet. The user pairs are the definition in every particular tweet.

Then, the last 2 features are std, and the age of the accounts. The age of accounts means how many days have these Twitter accounts been created.

IV. Result

For promoted tweets, accuracy 78%, precision 86%, recall 78%,F1 score 81.95%, AUC of ROC 79.3%.

For genuine tweets, accuracy 79%, precision 73%, recall 67%, F1 score 70.07%, AUC of ROC 76.73%.

Among all the features, top 3 dominating features in the classifier are:10th quantile of friends similarity,mean of followers similarity, and mean of retweeters account age.

V. Detour

In this section, all the works are unfinished and can not draw a confident conclusion. But all these works are based on the initial thoughts of this project. So the approaching steps are valuable.

As section I mentioned, the initial goal of this project is detecting ambiguous, unlabeled hype under the ground, but due to the difficulty and time limit we downgrade the goal into identifying promoted tweets.

In figure 4, it's a relation graph generated by *osome.iuni.iu.edu*, a useful social media analysis tool provided by Indiana University. By observing this figure, a user called *felberjosh* has many outlinks towards other users. Initially, this query was mean to find relationships between others and Donald Trump, since many experts thought his election result was based on hype. However, the result shows that the user called *Josh Felber* has a strong connection between himself and Donald Trump. Figure 4 and 5 show the query based on 2014. That's the time when Donald Trump was not decided to become a candidate of the president election. And Josh Felber did post some tweets supported Donald Trump. But those posts were supports on the business area. As we know, Donald Trump was famous as his personal success on business area at the beginning.

Figure 6 shows the homepage of Josh Felber, and intuitively saying, it's a hype account, observed by experience.

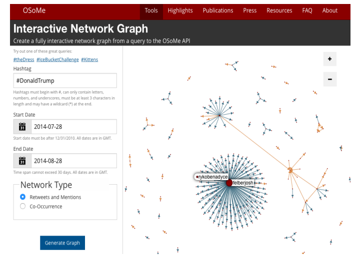

Fig. 4. Josh Felber relation graph

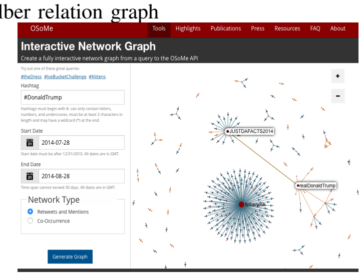

Fig. 5. Donald Trump relation graph

Fig. 6. An example of Josh Felber's Twitter homepage

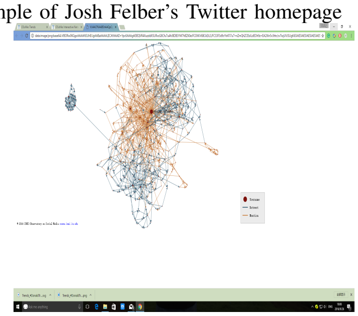

Fig. 7. Relation graph of Donald Trump

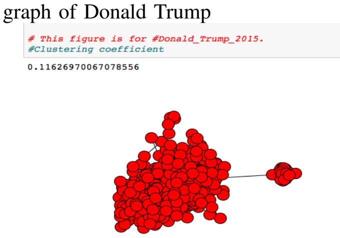

Fig. 8. Donald Trump cluster coefficient

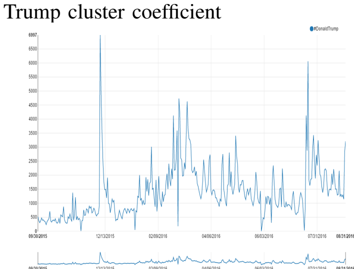

Fig. 9. Donald Trump Twitter trends

After finding out the relation graph, there're some analysis based on the graph. For example, figure 7 shows the retweet graph of Donald Trump. The interesting point is there's a small group of clusters on the top left corner. That's a group of tweets contributes to this topic, i.e.DonaldTrump. So I tried different approaches, including clustering coefficient and the shape of trends (Figure 8 and 9). However, Still no strong conclusions can be drawn.

FUTURE WORK

I think the initial thoughts of this project is on the right track, but I just lack enough knowledge to solve this problem. Maybe applying the same classification on this model will still provide the same confident result. However, there's no way to extract several clearly labeled hype tweets.

Thus, if there's no way to get confident input, the output will definitely full of errors.

This project will suspend for a while, until I have enough skills to make it more valuable.