Analyzing Spammers in Twitter User Network

Mananai Saengsuwan (@JustAProgrammer)

Motivation

- **Spams irritate users**
- **If we could identify spammers, we could block those accounts**
- **It’s a good ML and network analysis exercise**

Steps to do the analysis

1) Collect Twitter data
2) Implement a Spam classifier
3) Import Twitter data to the graph database
4) Analyze the data

The 2nd-4th steps will be done in python notebooks

Collecting Twitter data

- **There are a few resources cover this topic. The followings are samples:**
  - Bonzanini, Marco. “Users, Followers, and Communities on Twitter”. *Mastering Social Media Mining with Python*, Packt Publishing, 2016.
  - Saengsuwan, Mananai. “Building a network graph from Twitter data”. Web blog post. *Towards Data Science*. October 6, 2020.
  - Saengsuwan, Mananai. “Analyzing Twitter User Network”. Web blog post. *Towards Data Science*. June 23, 2022.

Collecting Twitter Data (Cont.)

Steps

- Identify 50 popular trends or hashtags
- For each trend, search for tweets
- For each tweet, identify user who writes or retweets it
- Finally, find friends of all those users

Collecting Twitter Data (Cont.)

- **The data we collected are:**
  - Statuses (Tweets)
  - Users
  - User-friends relationships
- **Also, we have labeled statuses (Spam or Ham) collected on different days for classification**
- **All data are in CSV format which could be imported to the graph database**

Spam classification

What we need

- **Labeled tweets file (training data)**
- **Unlabeled tweets file. The classifier would predict this data**

Output

- **Spam/Ham column in the tweets file**

Spam Classification(Cont.)

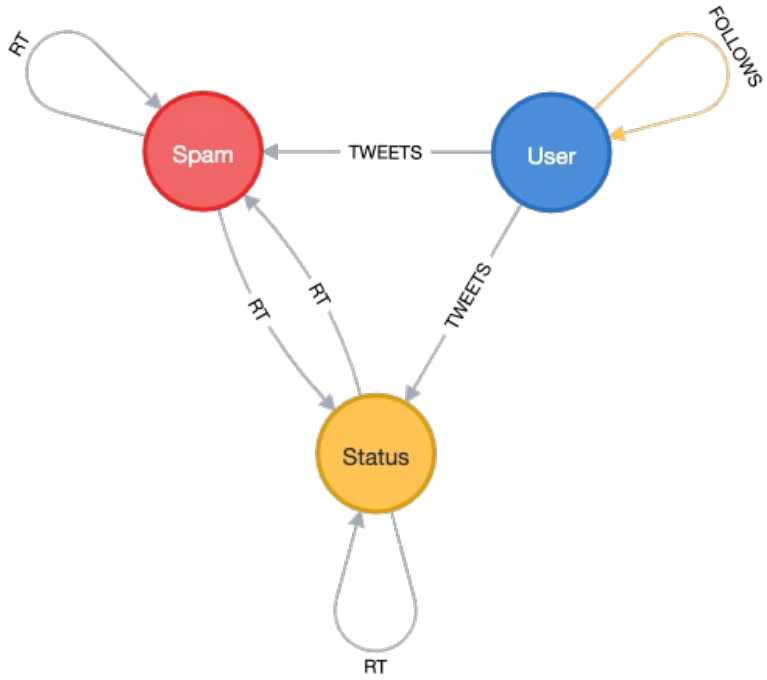
Steps:

1) **Read training tweet data file**
2) **Convert to feature vectors using TFID**
3) **Split the data into training and testing**
4) **Use Bayes as a classifier.**
5) **Fit it with the training data**
6) **Run the classifier with the testing data and check the scores (recall, precision)**
7) **Apply the classifier with the real tweet data will be used in the next step**

Importing data

- Copy the users, user_friends, and statuses files to the import directory
- Run import commands (Load CSV)
- Run a few Cypher commands to create relationships

Importing Data (Cont.)



Network analysis

- Visualization
- Triangle analysis
- Communities detection

Spammers Network Visualization using NX

Triangle analysis

- Three users follow each other. Two of them are spammer. The rest is a candidate
- To identify more spammers from existing ones.
- Limited success in our data.

Community detection

- Run Louvain community detection
- Identify communities with majorities are spammers. Assume all of them are spammers
- The community with a lot of spammer has very high relationship density. They follow each other a lot with in their group.
- In our data, we found one community. After a brief check, almost all are probably spammers

Conclusions

- From the visualization, spammers probably work together
- By using community detection, we could identify more spammers

Issue(s)

- **Spam classifier may misses some spams. We would not detect some spammers.**

Improvements

- **More labeled tweet data may improve classifier performance**
- **Twitter data collected here is only a tiny amount. But to collect more data:**
  - Need more time because of Twitter API rate limit
  - Bigger graph database consumes more resources (RAM and hard disk space)

Notable software library

In addition to several python libraries, I use the following Thai NLP software to tokenize text:

- **Wannaphong Phatthiyaphaibun, Korakot Chaovavanich, Charin Polpanumas, Arthit Suriyawongkul, Lalita Lowphansirikul, & Pattarawat Chormai. (2016, Jun 27). PyThaiNLP: Thai Natural Language Processing in Python. Zenodo. http://doi.org/10.5281/zenodo.3519354**

Github repository

For Twitter data collection:

- https://github.com/mananai/TwitterUserNetwork

For spam classification, data import, and spammer analysis:

- https://github.com/mananai/NODES2022-AnalyzingSpammers


Thank you