Weekly Report (June 8th - June 14th)

Xiutian Cui

June 14, 2015

1 DONE THIS WEEK

In this week I have finished the dataset parsing programming and started trying to detecting spammers by Machine Learning algorithms. The dataset I am using is combination datasets of [1] and [2]. [1] has been used as the graph structure of Twitter. And the suspended user list was depending on this dataset, by checking whether each user in this dataset is suspended now. [2] has provided the tweets and detail profile of users. Also I checked the changing from [1] to [2] during these two years.

I built one website (http://twitter.jekycui.net) to describe what I am doing by details. And the code of parsing these two dataset were provided on this website as well.

I have started to try some classification on the merged dataset. I have merged these two dataset and converted it to an ARFF file, which can be imported by Weka. I tried some classifiers in Weka, but the results are not good (see Figure 1.1). The accuracy of JRip, which are the best classifier in [3], is only 57%.

2 PLAN NEXT WEEK

I will try to figure out why the classification result is so low and continue writing proposal slides.

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances        3471               57.0982 %
Incorrectly Classified Instances      2608               42.9018 %
Kappa statistic                          0.142
Mean absolute error                      0.4846
Root mean squared error                  0.495
Relative absolute error                 96.9177 %
Root relative squared error             99.0022 %
Total Number of Instances             6079

=== Detailed Accuracy By Class ===

                TP Rate   FP Rate   Precision   Recall   F-Measure   ROC Area   Class
                0.57      0.428     0.572       0.57     0.571       0.582      spammer
                0.572     0.43      0.57        0.572    0.571       0.582      non-spammer
Weighted Avg.   0.571     0.429     0.571       0.571    0.571       0.582

=== Confusion Matrix ===

    a    b    <-- classified as
 1735 1310 |   a = spammer
 1298 1736 |   b = non-spammer
```

Figure 1.1: Result of JRip in Weka

REFERENCES

[1] KWAK, H., LEE, C., PARK, H., AND MOON, S. What is Twitter, a social network or a news media? In *WWW '10: Proceedings of the 19th international conference on World wide web* (New York, NY, USA, 2010), ACM, pp. 591–600.

[2] LI, R., WANG, S., DENG, H., WANG, R., AND CHANG, K. C.-C. Towards social user profiling: unified and discriminative influence model for inferring home locations. In *KDD* (2012), pp. 1023–1031.

[3] MOH, T.-S., AND MURMANN, A. J. Can you judge a man by his friends?-enhancing spammer detection on the twitter microblogging platform using friends and followers. In *Information Systems, Technology and Management*. Springer, 2010, pp. 210–220.