CS670 Project Proposal
Fraudulent Lockstep Behavior Detection

Team Name: **End of Fraud**

Team Members (alphabetically ordered):

Majid Alfifi, Parisa Kaghazgaran, Xing Zhao

1 Introduction

How can we detect if a politician has purchased fake followers on Twitter or if a product's reviews on Amazon are not genuine?

A common method has been to represent *users* and *items* as a matrix where in the simplest case a cell can take on a binary value of 1 or 0 indicating whether there is a relationship between the corresponding user and item or not. The problem can then be transformed to finding dense regions in this matrix [1]. Moreover, this method has been lately extended from matrix to tensor representation to incorporate more dimensions from the domain such as timestamp, Twitter followers count, or number of stars of an Amazon product [3] [4]. Extraordinary dense blocks in the tensor correspond to groups of users with lockstep behaviors both in the items they relate to and along the additional dimensions (for example, multiple users reviewing the same products at the exact same time). A scalable MapReduce-based implementation **D-Cube**[1] was also provided by [4] which we intend to evaluate and build on in this project.

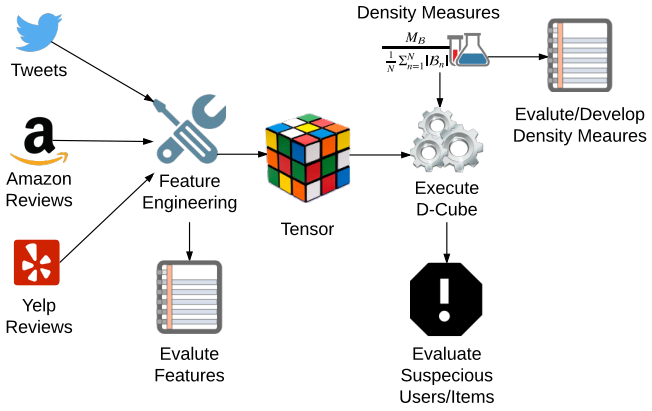

Figure 1. Project plan: *Feature engineering* will help find the most useful dimensions to use in the Tensor to enable better suspicious blocks detection. *Density measures* experiments will evaluate different existing measures in addition to custom measures that might be suitable for our own datasets.

2 Project Goals

As illustrated in Figure 1, we intend to use an existing MapReduce-based implementation of the D-Cube algorithm (Section 3.3 in [4]) on our own datasets with two main goals:

1. **Feature engineering**: we will explore the effectiveness of different dimensions in each of our datasets in detecting fraudulent lockstep behavior; hopefully also informing our own research.

2. **Empirical study of density measures**: we aim to experiment and build on several density measures defined in [4] (Section 2.2) and its foundations previously laid in [2] and potentially propose our own flexible density measures that take into account weights of different features in different datasets. For example, In Amazon dataset, is the temporal dimension more informative than product rating?

3 Datasets

We intend to apply and experiment with the D-Cube algorithm on the following datasets:

- **Twitter dataset:** 9 billion tweets of which 2 billion are from eventually suspended users. This dataset was obtained through an industry contact.

[1] https://github.com/kijungs/dcube

Questions: Can we identify tweets/users tampering with hashtags in efforts to promote/undermine discussions in those hashtags? How do those suspicious tweets temporally differ from ordinary tweets (timestamp dimension)? Can we identify users hired to attack other users by continuously replying to their tweets casting doubt on their cause? etc.

- **Amazon dataset:** About 500,000 Amazon reviews of which 21,000 are from crowdturfing products and the rest are from activity history of reviewers who wrote a review under these products. This dataset was collected from Amazon.com

  Questions: Can we find the hidden relation among reviewers in Amazon who aim to promote a product by writing fake reviews using additional dimensions such as temporal, rating and so on.

- **Yelp dataset:** 4.1M reviews by 1M users for 144K businesses. This dataset is made available for researchers by Yelp[2].

  Questions: Can we identify fraud users (e.g. who gave the reviews for multiples businesses in a same day) who were hired to give fake reviews to a business? What are their behaviors (e.g. their average ratings) in their reviews? What kind of features (e.g. geolocation, business type, etc.) do businesses reviewed by fraud users have?

4 Evaluation

The D-Cube paper [4] evaluates performance of the algorithm on synthetic data (i.e. by injecting random dense blocks in the tensor). However, we intent to provide a more realistic evaluation of the quality of each dense block (or blocks) found by D-Cube as follows:

- **Twitter dataset:** How many of the suspicious users found were eventually suspended or deleted by Twitter? What is the size of the intersection between top hashtags used by suspended users and hashtags identified by D-Cube?

- **Amazon dataset:** How many of the suspicious products found were actually for products known to have explicitly recruited reviewers on crowdsourcing websites?

- **Yelp dataset:** How many of the suspicious reviews found were filtered by Yelp?

[2]https://www.yelp.com/dataset

In addition, we will sample few hundred users and items from each block we find and investigate them manually using our own domain knowledge of how fraudulent behavior in each domain looks like.

5 Tools/Resources

In addition to all the tools and methods we learned in the course, we will use a local Hadoop cluster to run the D-Cube algorithm.

6 Project Outcome

- Evaluation of the accuracy of dense suspicious blocks found in each dataset.

- Evaluation of different dimensions useful for each dataset.

- Evaluation of different density measures for each dataset.

- (Optional) New density measures that are more suitable to the nature of the datasets we study.

References

[1] B. Hooi, H. A. Song, A. Beutel, N. Shah, K. Shin, and C. Faloutsos. Fraudar: Bounding graph fraud in the face of camouflage. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 895–904. ACM, 2016.

[2] M. Jiang, A. Beutel, P. Cui, B. Hooi, S. Yang, and C. Faloutsos. A general suspiciousness metric for dense blocks in multimodal data. In *Data Mining (ICDM), 2015 IEEE International Conference on*, pages 781–786. IEEE, 2015.

[3] K. Shin, B. Hooi, and C. Faloutsos. M-zoom: Fast dense-block detection in tensors with quality guarantees. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 264–280. Springer, 2016.

[4] K. Shin, B. Hooi, J. Kim, and C. Faloutsos. D-cube: Dense-block detection in terabyte-scale tensors. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, pages 681–689. ACM, 2017.