

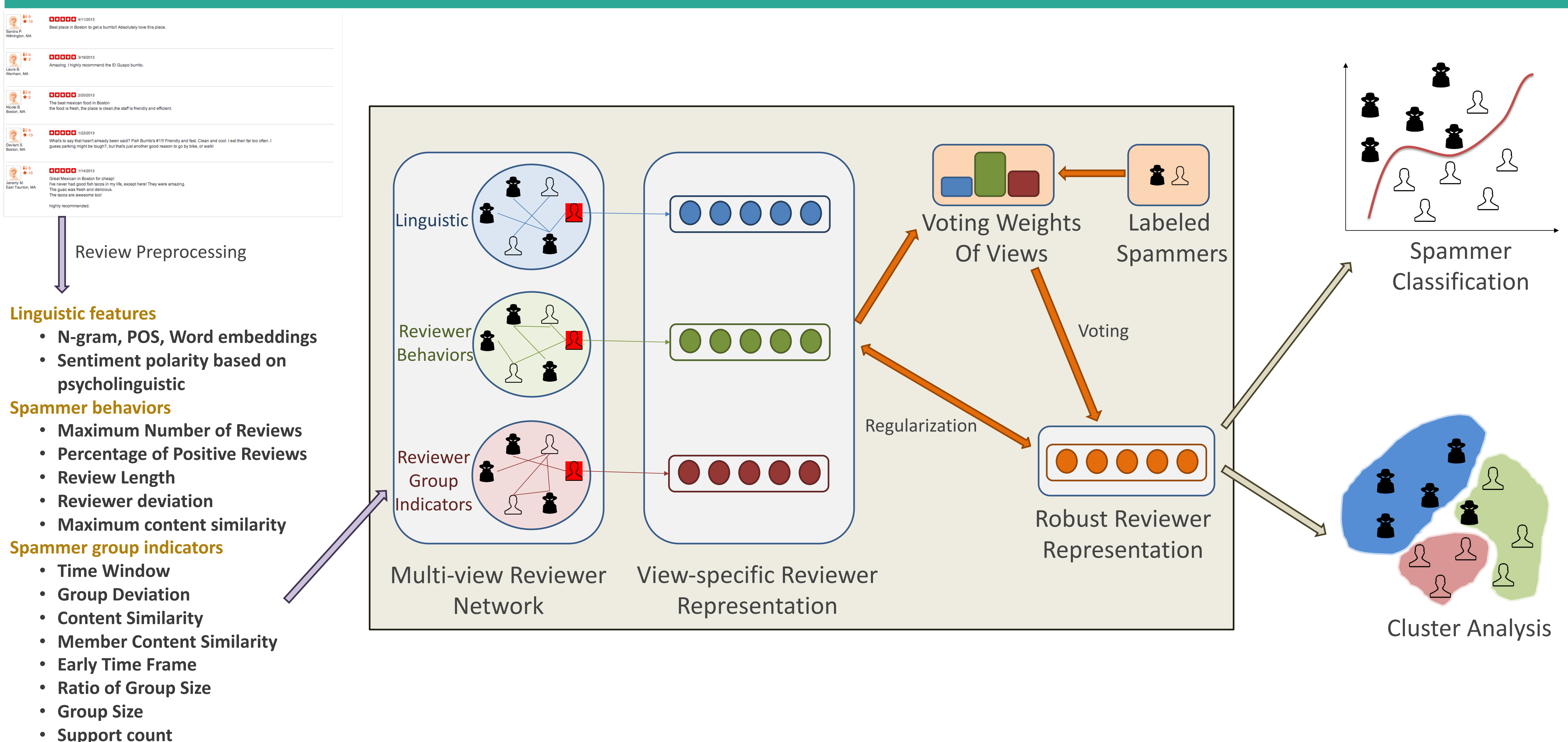
## Leveraging Multi-View Reviewer Representation for Spammer Detection in Online Reviews

Shih-Feng Yang (yang798@purdue.edu), Julia Taylor Rayz (jtaylor1@purdue.edu)

### Background

This poster presents a proposal and preliminary results by applying multi-view network representation learning for spammer detection on online review datasets collected from Amazon and Yelp. Previous research has focused on the textual characteristics of reviews, spammer behaviors, and the common patterns in a spammer group. A recent trend of utilizing network inference on spam detection enables us to jointly consider spam reviews, spammers, and spammer groups as heterogenous networks. However, several gaps exist in the literature on whether reviewer networks were properly designed for spam detection. First, previous studies considered only first-order proximity but did not include second-order proximity (Tang et al., 2015) when inferring the structure of reviewer networks. Second, in real world, reviewers can form various network views with respect to different features. Previous studies treated reviewer networks as single-view networks rather than multi-view networks. In this study, we adopt MVE (Qu et al., 2017) to learn robust reviewer embedding from multi-view reviewer network. We use the learned reviewer embedding as reviewer's features and investigate its effectiveness on spammer classification. Moreover, we conduct a cluster analysis on reviewer embeddings to evaluate the correlation between a reviewer's spam tendency and its cluster.

### Solution



### Preliminary Results

**Three-View Reviewer Network regarding Spamming Behaviors (Mukherjee et al., 2013)**

- Maximum Number of Reviews (MNR)
- Percentage of Positive Reviews (PR)
- Review Length (RL)

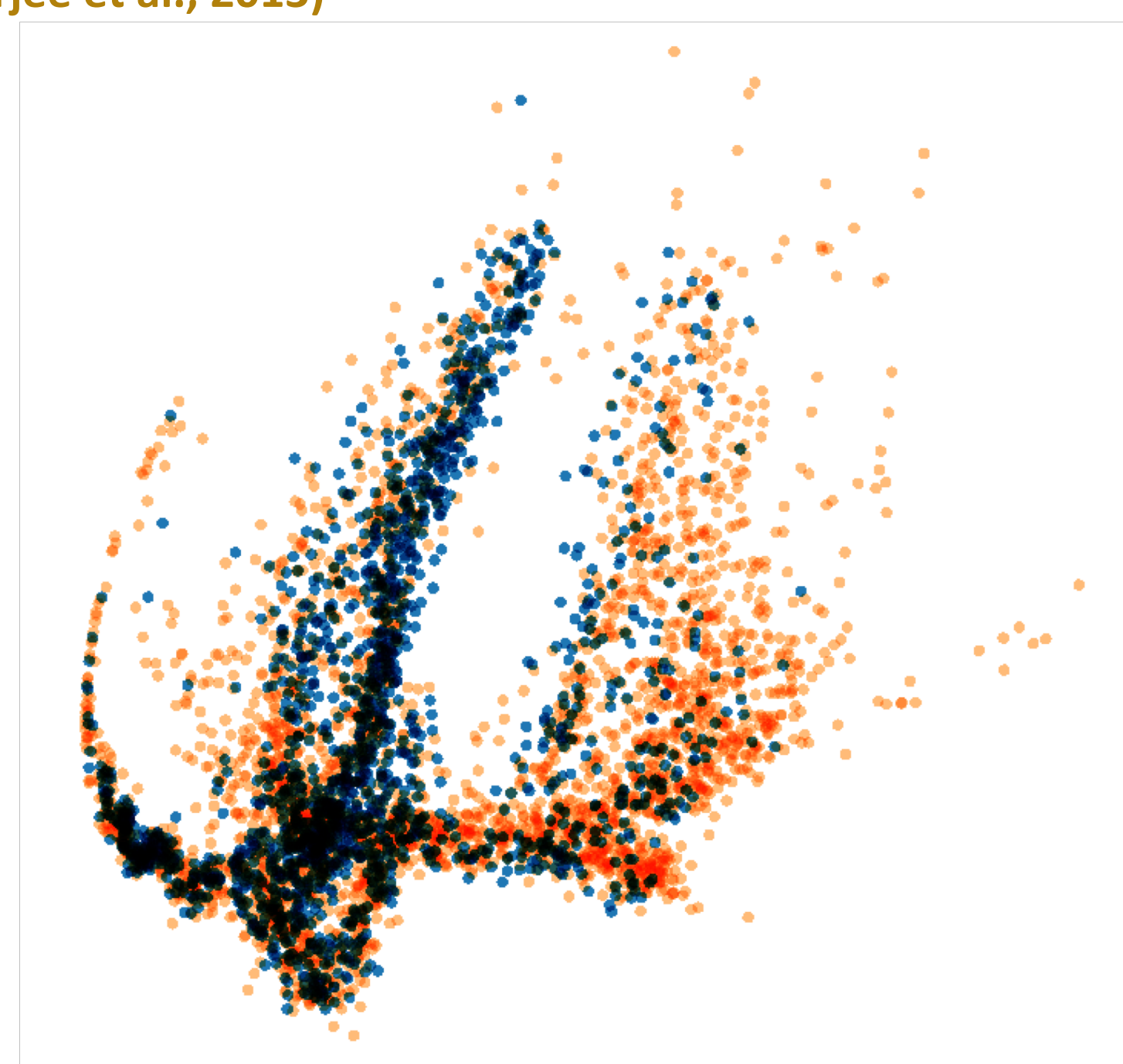
**YelpNYC Dataset (Rayana and Akoglu, 2015)**

- Training: 4000 reviewers; Testing: 1000 reviewers.

TRAINING RESULTS	PRECISION	RECALL	F1-SCORE
Non-Spammer	57%	39%	46%
Spammer	70%	83%	76%
micro avg	67%	67%	67%
macro avg	64%	61%	61%
weighted avg	66%	67%	65%

TESTING RESULTS	PRECISION	RECALL	F1-SCORE
Non-Spammer	54%	41%	46%
Spammer	70%	80%	75%
micro avg	66%	66%	66%
macro avg	62%	60%	61%
weighted avg	64%	66%	65%

Spammer Classification



2-D PCA of Reviewer Embeddings  
● spammer ● non-spammer

### References

- Mukherjee, A., Venkataraman, V., Liu, B., & Glance, N. (2013, June). What yelp fake review filter might be doing?. In Seventh international AAAI conference on weblogs and social media.
- Qu, M., Tang, J., Shang, J., Ren, X., Zhang, M., & Han, J. (2017, November). An attention-based collaboration framework for multi-view network representation learning. In Proceedings of the 2017 ACM on Conference on Information and Knowledge Management (pp. 1767-1776). ACM.
- Rayana, S., & Akoglu, L. (2015, August). Collective opinion spam detection: Bridging review networks and metadata. In Proceedings of the 21th acm sigkdd international conference on knowledge discovery and data mining (pp. 985-994). ACM.
- Tang, J., Qu, M., Wang, M., Zhang, M., Yan, J., & Mei, Q. (2015, May). Line: Large-scale information network embedding. In Proceedings of the 24th international conference on world wide web (pp. 1067-1077). International World Wide Web Conferences Steering Committee.