

Beyond News Contents: The Role of Social Context for Fake News Detection

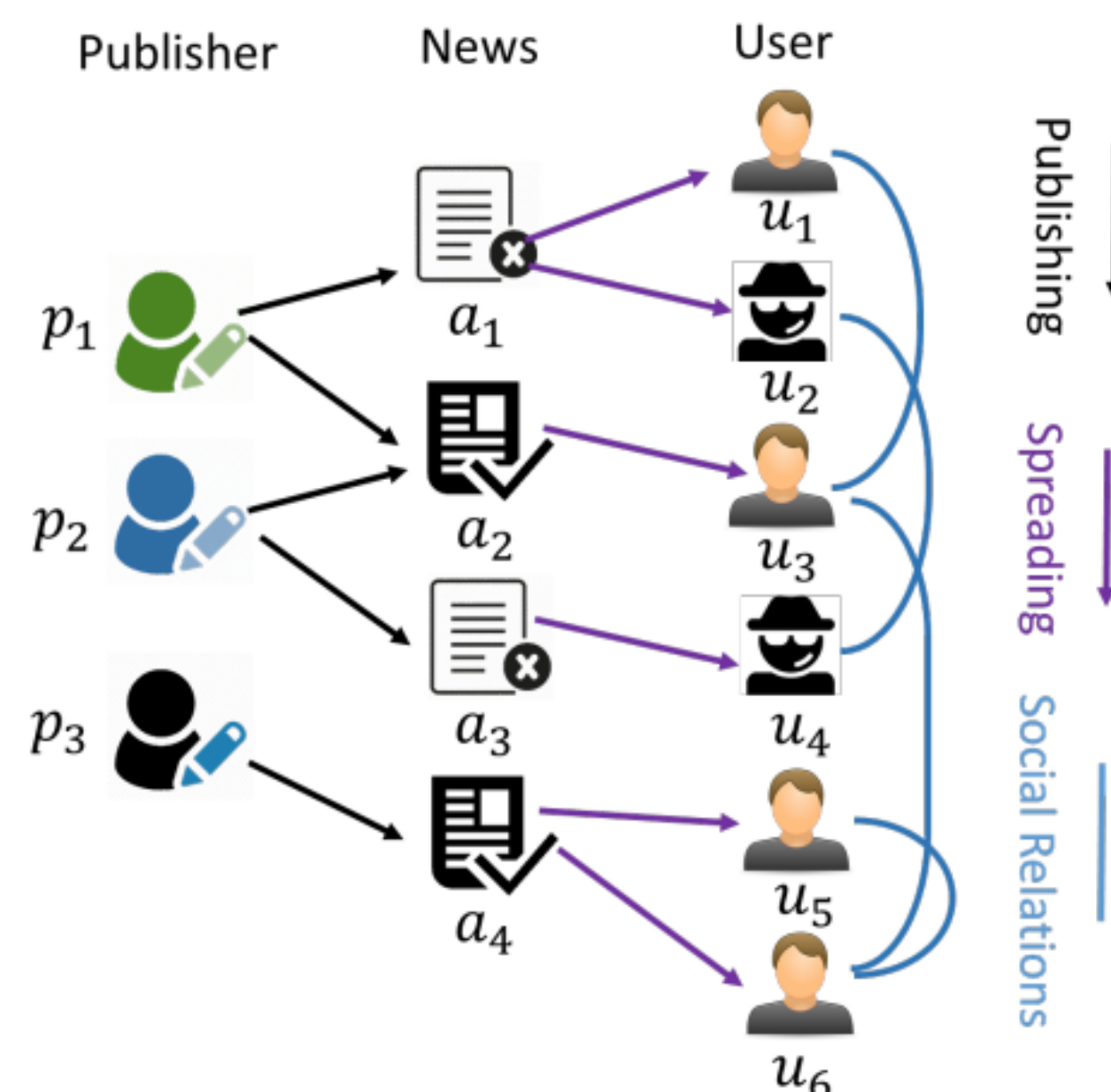
Kai Shu¹, Suhang Wang², Huan Liu¹

¹Arizona State University, ²Penn State University

{kai.shu, huan.liu}@asu.edu, szw494@psu.edu

Introduction

- Fake news is widely spread on social media and has detrimental societal effects
- Would social context information be helpful in making fake news detection?
- A typical news dissemination ecosystem:
 - Entities: publisher p , news a , and social media user u
 - Relations: publishing, spreading, social relations



- Contributions:** (i) A principled way to model tri-relationship among publishers, news pieces, and users simultaneously; (ii) A novel framework TriFN, which exploits both user-news interactions and publisher-news relations for learning news feature representations to predict fake news

The Proposed Framework - TriFN

- Tri-Relationship Embedding Framework

- News content embedding

$$\min_{\mathbf{D}, \mathbf{V} \geq 0} \|\mathbf{X} - \mathbf{D}\mathbf{V}^T\|_F^2 + \lambda(\|\mathbf{D}\|_F^2 + \|\mathbf{V}\|_F^2)$$

- Publisher-News Relation Embedding

$$\min \|\bar{\mathbf{B}}\mathbf{D}\mathbf{Q} - \mathbf{o}\|_2^2 + \lambda\|\mathbf{Q}\|_2^2$$

- User Social Relation Embedding

$$\min_{\mathbf{U}, \mathbf{T} \geq 0} \|\mathbf{Y} \odot (\mathbf{A} - \mathbf{U}\mathbf{T}\mathbf{U}^T)\|_F^2 + \lambda(\|\mathbf{U}\|_F^2 + \|\mathbf{T}\|_F^2)$$

- User-News Interaction Embedding

$$\min \underbrace{\sum_{i=1}^m \sum_{j=1}^r w_{ij} c_i \left(1 - \frac{1 + y_{Lj}}{2}\right) \|u_i - \mathbf{D}_{Lj}\|_2^2}_{\text{True news}} + \underbrace{\sum_{i=1}^m \sum_{j=1}^r w_{ij} (1 - c_i) \left(\frac{1 + y_{Lj}}{2}\right) \|u_i - \mathbf{D}_{Lj}\|_2^2}_{\text{Fake news}}$$

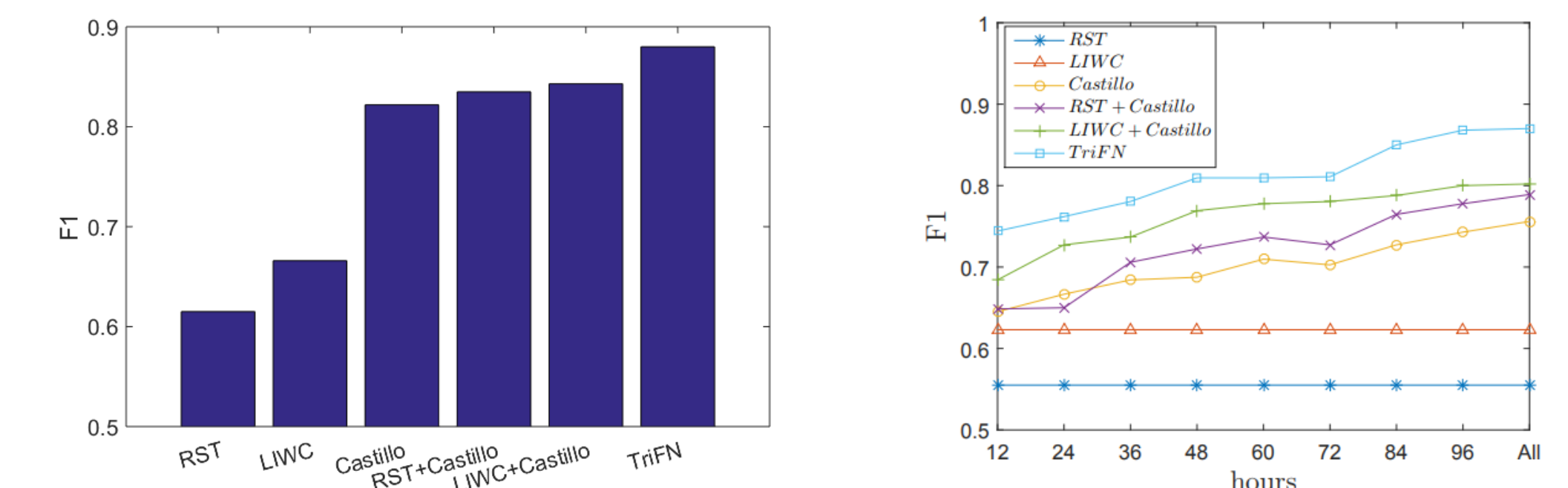
- Objective function of TriFN

$$\min_{\mathbf{D}, \mathbf{U}, \mathbf{V}, \mathbf{T} \geq 0, p, q} \|\mathbf{X} - \mathbf{D}\mathbf{V}^T\|_F^2 + \alpha \|\mathbf{Y} \odot (\mathbf{A} - \mathbf{U}\mathbf{T}\mathbf{U}^T)\|_F^2 + \beta \text{tr}(\mathbf{H}^T \mathbf{L}\mathbf{H}) + \gamma \|\mathbf{o} \odot (\bar{\mathbf{B}}\mathbf{D}\mathbf{q} - \mathbf{o})\|_2^2 + \eta \|\mathbf{D}_{LP} - y_L\|_2^2 + \lambda R$$

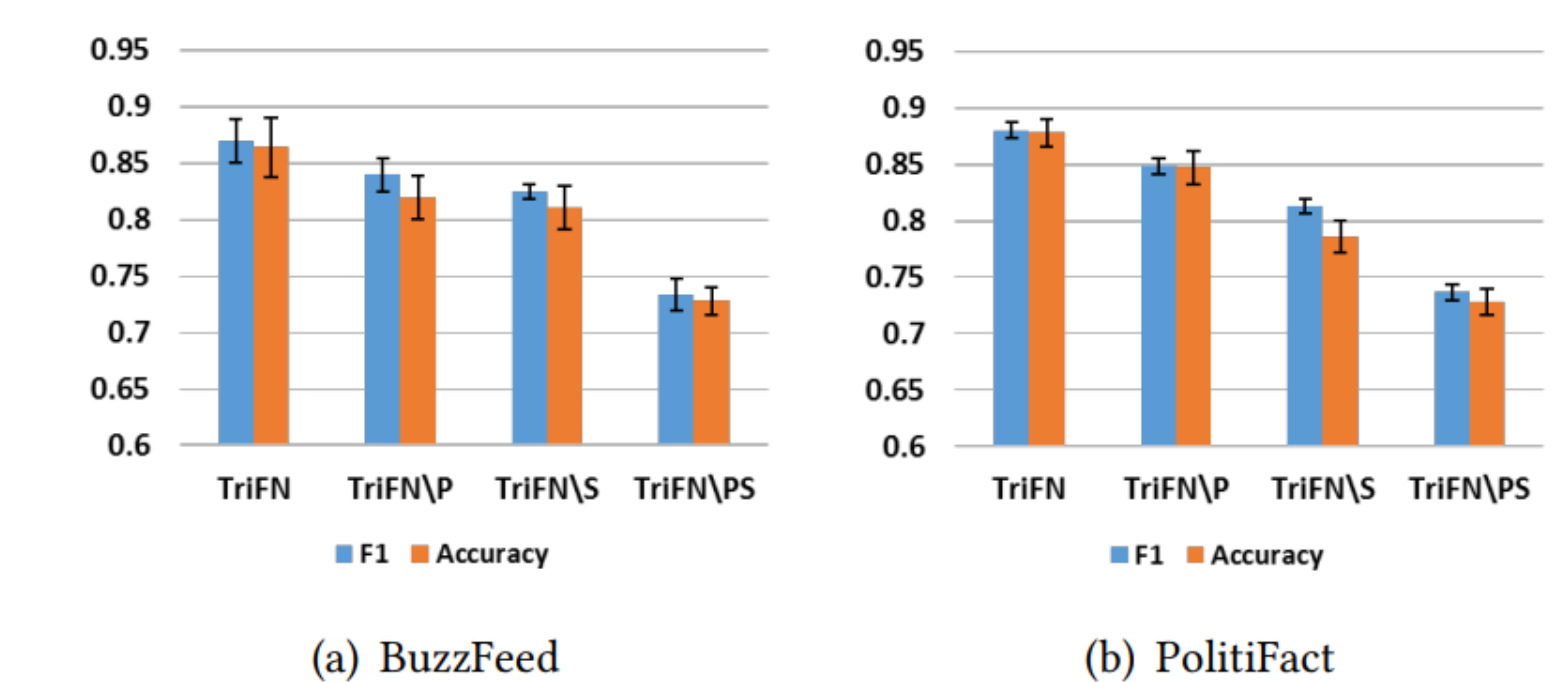
- We jointly combine news content embedding and social context embedding for fake news detection

Experimental Results

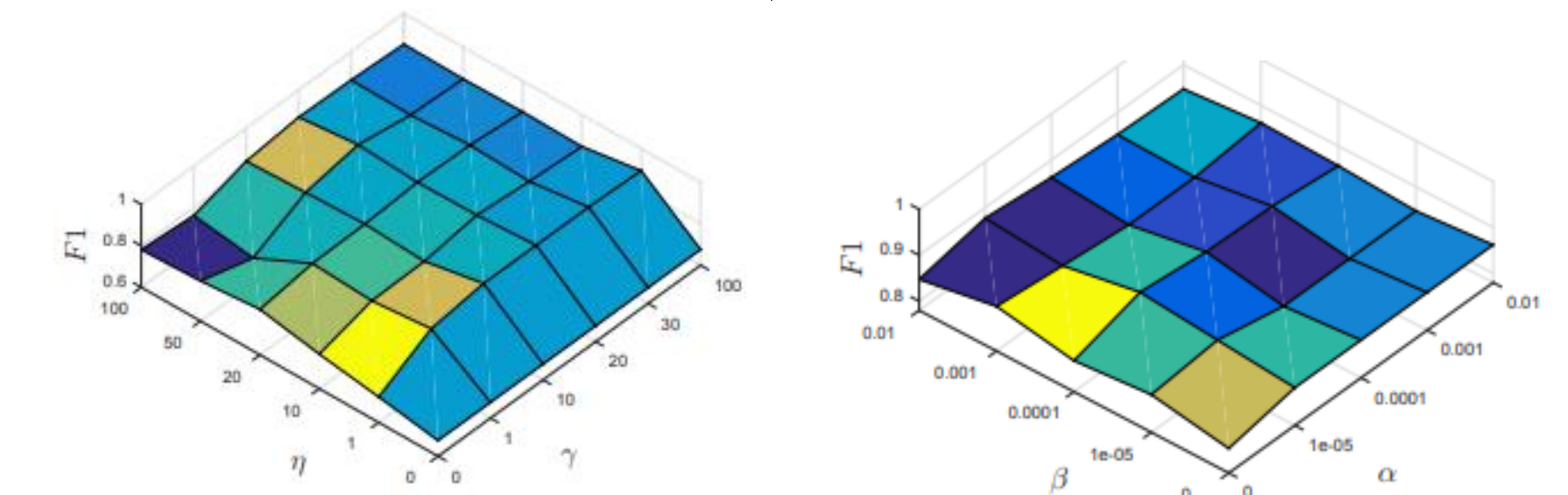
- Fake news (early) detection performance



- Impact analysis of users and publishers



- Parameter Sensitivity (on Politifact dataset)



Future Work

- (i) Extend TriFN to scenario with limited explicit user propagation paths are given;
- (ii) Explore social context for unsupervised fake news detection.

This material is based upon work supported by, or in part by, NSF #1614576, and ONR N00014-16-1-2257.