

Paper Summary

David Miller
CIS 5930: Social Network Mining

February 19, 2018

The purpose of this paper is to present a method that can be used for detecting rumors or credible new users from social media. Mathematically we let $\mathbf{D} = \{\mathbf{d}_1, \dots, \mathbf{d}_m\} \in \mathbb{R}^{m \times n}$ be a data matrix with row $\mathbf{d}_1 \in \mathbb{R}^n$ being a data instance and column $\mathbf{f}_i \in \mathbb{R}^m$ a vector of the i -th feature. We also let $\mathbf{y} \in \{-1, 1\}^m$ be a label vector such that $y_i = 1$ if i refers to a rumor and -1 otherwise. Given \mathbf{D} and \mathbf{y} , the main goal is to learn a predictor to accurately classify rumors and non-rumors based on social media input [1]. From [2] the matrix factorization problem

$$\min_{\mathbf{U}, \mathbf{V}} \frac{1}{2} \|\mathbf{D} - \mathbf{U}\mathbf{V}^T\|_F^2, \quad s.t. \mathbf{U} \in \{0, 1\}^{m \times k}, \mathbf{U}\mathbf{1} = \mathbf{1}, \quad (1)$$

where $\mathbf{1}$ is a vector of 1s and $\|\cdot\|_F$ is the Frobenius norm, is known to solve latent representations. $\mathbf{U} \in \mathbb{R}^{m \times k}$ is the low-rank representation of users and $\mathbf{V} \in \mathbb{R}^{n \times k}$ is the low-rank representation of features. This factorization separates data from feature by k latent factors. After manipulation, the paper ends up at an equation (omitted for this review) that has three parameters \mathbf{U} , \mathbf{V} , \mathbf{w} which becomes the objective function to minimize. Since the objective function is not convex with respect to all three parameters, but is to each individually, stochastic gradient descent is applied to a parameter while fixing the other two iteratively. The rumor classifier is obtained once given a fixed \mathbf{U} and \mathbf{V} . the algorithm is

Algorithm 1 Early Detection of Emerging Rumors

Input: Data matrix \mathbf{D} , label vector \mathbf{y} , maximal number of iterations I

Output: \mathbf{U} , \mathbf{V} , \mathbf{w}

```
1:   Generate  $\mathbf{U}$ ,  $\mathbf{V}$ , and  $\mathbf{w}$  randomly
2:   for  $i = 1$  to  $I$  do
3:       Update  $\mathbf{U}$ 
4:       Update  $\mathbf{V}$ 
5:       Update  $\mathbf{w}$ 
6:       if convergence then
7:           break
8:       end if
9:   end for
10:  return  $\mathbf{U}$ ,  $\mathbf{V}$ ,  $\mathbf{w}$ 
```

From this rumors and non-rumors can be classified from the returned parameters \mathbf{U} , \mathbf{V} , \mathbf{w} . In terms of future direction, the algorithm can use other sources of information than just social media. The paper talks about cross-modal information, but another thing I had in mind was geo-location. Sources near a rumors point of interest are more likely to be trustworthy than sources far from the rumors point of interest.

Three strengths I found with the paper are

1. The accuracy of detecting rumors outperforms all its competitors it was tested against (Pooling, Elastic Net, KM SVM, FE LL, LK RBF) with somewhat significant margins.
2. Rumor detection algorithm lends itself to countless applications, and the paper did a fantastic job of outlining its potential uses. This is very good for people who are reading the paper and have commercial endeavors in mind.
3. The analysis shows that parts of the algorithm can be done in $\mathcal{O}(k^3)$ time, where $k \ll n$. This means that the algorithm is dependent of the amount of clusters, or categories.

Three weaknesses I found with the paper are

1. There is no mention of what system the algorithm was tested on. This means the entire algorithm may need warehouse scale computers to be driven.
2. The algorithm relies on a huge data size, making in impractical for rumors/non-rumors of little importance (not many people talking about it).
3. The objective function can not used gradient descent directly, but rather needs to fix certain parameters while applying gradient descent to the dynamic parameter. This can lead to non-convergence.

Questions for the reader

1. As of now I have no questions but I am sure I will have some during the presentation.

References

- [1] Liang Wu, Jundong Li, Xia Hu, Huan Liu *Gleaning Wisdom from the Past: Early Detection of Emerging Rumors in Social Media*, KDD17, August 13-17, 2017, Halifax, NS, Canada.
- [2] Jundong Li, Xia Hu, Liang Wu, Huan Liu *Robust Unsupervised Feature Selection on Networked Data*, KDD17, August 13-17, 2017, Halifax, NS, Canada.