2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining

Semi-Supervised Learning and Graph Neural Networks for Fake News Detection

Adrien Benamira*, Benjamin Devillers*, Etienne Lesot*, Ayush K. Ray*, Manal Saadi*, Fragkiskos D. Malliaros*†
*CentraleSupélec, University of Paris-Saclay, France
Email: {adrien.benamira, benjamin.devillers, etienne.lesot, manal.saadi}@supelec.fr, ayush.rai2512@student-cs.fr
†Inria Saclay, France
Email: fragkiskos.malliaros@centralesupelec.fr

***Abstract*—Social networks have become the main platforms for information dissemination. Nevertheless, due to the increasing number of users, social media platforms tend to be highly vulnerable to the propagation of disinformation – making the detection of fake news a challenging task. In this work, we focus on content-based methods for detecting fake news – casting the problem to a binary text classification one (an article corresponds to either fake news or not). In particular, our work proposes a graph-based semi-supervised fake news detection method based on graph neural networks. The experimental results indicate that the proposed methodology achieves better performance compared to traditional classification techniques, especially when trained on limited number of labeled articles[1].**

***Index Terms*—Fake news detection, Semi-supervised learning, Graph neural networks**

I. Introduction

Social media have become the main platforms for information sharing and news consumption for various reasons. Firstly, it is often a faster and cheaper way to access news on social media compared to more traditional platforms. Furthermore, commenting, sharing and discussing with other readers is an easy way to express opinions and increases the level of participation and interaction of individuals. Nevertheless, the ease with which real-time information disseminates to a large audience accompanied with the engagement of individuals to online social media platforms, has also lead to the spread of misinformation, widely known as *fake news* [8]. Fake news take advantage of the *echo chambers* phenomenon, amplified by social networks; people tend to follow and share mainly information they believe in or what their friends share and like.

Triggered by the societal impact of misinformation, there is currently an intense research effort from the scientific community to develop algorithmic techniques for fake news detection. The core of these techniques relies on machine learning methods – trying to analyze and understand how the content of fake news differs from that of real ones, as well as how users engage with and propagate misinformation within social networks [8], [10].

In order to tackle the fact that labels are often very limited and sparse, in our approach we opt for *semi-supervised* content-based detection methods. In particular, we propose a graph-based semi-supervised fake news detection framework, building upon network representation learning techniques [3]. Our intuition is that, graphs are expressive models that are able to capture contextual dependencies among articles, alleviating the label scarcity constraint [2]. On a high level, our framework is composed of three components: i) embedding of articles in the Euclidean space; ii) construction of an article similarity graph; iii) inference of missing labels using graph learning techniques. The main contributions of this paper are summarized as follow:

- We use word embeddings to obtain latent representations of news articles in a lower dimensional Euclidean space. Then, we capture contextual similarities among articles via a graph-based representation scheme.
- We cast the fake news detection problem to a semi-supervised graph learning task, leveraging Graph Neural Network architectures that are able to perform well on limited labeled data.
- We perform a preliminary evaluation of our methodology on a real fake news dataset, demonstrating that the proposed methodology outperforms previous content-based approaches, requiring fewer labeled articles.

II. Proposed Approach

A schematic representation of our approach is depicted in Fig. 1. Considering a collection of M articles, our graph-based method is composed by the following three components:

- *Embedding of articles*: Here we are considering a vector representation of an article. Our approach is based on pre-trained GloVe word embeddings, computing the mean vector of the words appearing within an article.
- *Graph construction*: The second step concerns the construction of similarity graph among articles [2]. In par-

[1]Our code is publicly available at: https://github.com/bdvllrs/misinformation-detection-tensor-embeddings.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

ASONAM '19, August 27-30, 2019, Vancouver, Canada
© 2019Association for Computing Machinery.
ACM ISBN 978-1-4503-6868-1/19/08
http://dx.doi.org/10.1145/3341161.3342958

Authorized licensed use limited to: Hebei University of Technology. Downloaded on January 04,2021 at 06:08:38 UTC from IEEE Xplore. Restrictions apply.

TABLE I
CLASSIFICATION ACCURACY WITH DIFFERENT RATIOS OF LABELED DATA USED FOR TRAINING WITH $k = 4$.

Methods	Accuracy (in %)				
	2 % labeled data	*5 % labeled data*	*10 % labeled data*	*15 % labeled data*	20 % labeled data
Guacho et al. [2]	56.65 ± 9.67	63.60 ± 7.52	70.95 ± 5.28	74.05 ± 3.80	79.8 ± 3.10
SVM	63.55 ± 5.73	66.55 ± 7.14	75.05 ± 5.20	76.05 ± 4.80	78.90 ± 5.16
Random Forest	60.25 ± 10.02	69.05 ± 3.33	76.65 ± 3.48	**83.55 ± 5.06**	84.70 ± 2.48
AGNN	70.45± 5.39	72.00 ± 8.05	78.70 ±3.54	83.35 ± 1.74	84.25 ± 3.51
GCN	**72.04 ± 6.00**	**77.35 ± 3.72**	**79.85 ± 3.41**	82.35 ± 2.44	**84.94 ± 2.30**

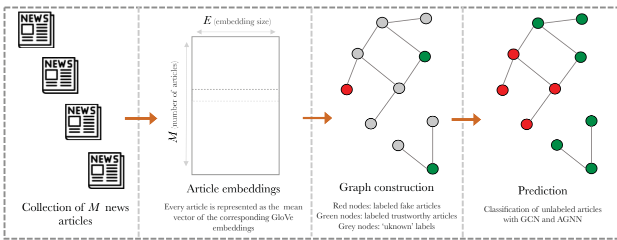

Fig. 1. Illustration of the proposed approach: M denotes the number of articles (real and fake) and E is the dimension of our GloVe embeddings (in our case, $M = 150$, $E = 100$). Finally, we use $k = 4$ nearest neighbours to build the graph.

ticular, for each article (i.e., node in the graph), we look for the k–nearest neighbours based on the Euclidean distances in the embedding space.

- *Classification*: For the classification task over the similarity graph between articles, we use two graph neural network methods: Graph Convolutional Networks (GCN) [5] and Attention Graph Neural Network (AGNN) [9].

III. EXPERIMENTAL EVALUATION

Experimental Set-up. In our empirical evaluation, we have compared our graph neural network-based methods against the approach by Guacho et al. [2], which follows a similar framework. In this baseline method, the embedding of articles is obtained using CP/PARAFAC tensor decomposition on the binary co-occurrence matrix between all articles, and the classification is performed using the Fast Belief Propagation (FaBP) algorithm [6]. We have also compared against a traditional approach, which consists of bag-of-words textual features and learning with SVMs and Random Forest classification models. For the neural graph networks, we use 4 layers, 4 neighbours, 16 hidden units, a learning rate of 0.01 and a weight decay of $5e-4$. We train our graph neural networks during 1000 epochs and we keep the one which has the best accuracy on the test (unlabeled) data. We evaluate our method on a recent dataset [4], which is comprised of 150 labeled articles, 75 of those are fake news and 75 real. We pick the labeled articles at random and average the results over 20 independent experiments.

Results. Table I gives the classification accuracy of the different methods, varying the amount of labeled data used for training (ranging from 2% up to 20%). As we can observe, the proposed graph neural network approaches achieve a performance improvement of up to 3% with only 10% of the labeled data, while being more stable and reducing the standard deviation of the results. Furthermore, our AGNN and GCN methods are computationally faster when it comes to evaluate a new article. In a more extended version of this article [1], we also examine the influence of different article embedding techniques.

IV. CONCLUSION AND FURTHER WORK

In this paper, we have focused on content-based misinformation detection, assuming that we have a limited amount of labeled articles. The preliminary experimental results suggested that building a simple nearest-neighbour graph among articles based on word embedding similarities accompanied by graph neural networks for classification, gives qualitatively good results — providing the basis for semi-supervised content-based detection methods. We are currently working to further extend our study by considering more baseline methods and testing the performance on bigger as well as multi-labeled fake news datasets.

REFERENCES

[1] Benamira, A., Devillers, B, Lesot, E., Ray, A.K., Saadi, M., and Malliaros, F.D. (2019). Semi-supervised learning and graph neural networks for fake news detection. *Technical Report*.

[2] Guacho, G. B., Abdali, S., Shah, N., and Papalexakis, E. E. (2018). Semi-supervised Content-based Detection of Misinformation via Tensor Embeddings. In *ASONAM*, pp. 322-325.

[3] W. L. Hamilton, Ying, R. and Leskovec, J. (2017). Representation learning on graphs: Methods and applications. *IEEE Data Eng. Bull.*, 40(3):52–74.

[4] Horne, B. D., and Adali, S. (2017). This just in: fake news packs a lot in title, uses simpler, repetitive content in text body, more similar to satire than real news. In *ICWSM* Workshops.

[5] Kipf, T. N., and Welling, M. (2016). Semi-supervised classification with graph convolutional networks. In *ICLR*.

[6] Koutra, D., Ke, T. Y., Kang, U., Chau, D. H. P., Pao, H. K. K., Faloutsos, C. (2011). Unifying guilt-by-association approaches: Theorems and fast algorithms. In *ECML PKDD*, pp. 245-260.

[7] Scarselli, F., Gori, M., Tsoi, A. C., Hagenbuchner, M., Monfardini, G. (2009). The graph neural network model. *IEEE Transactions on Neural Networks*, 20(1), 61-80.

[8] K. Sharma, F. Qian, H. Jiang, N. Ruchansky, M. Zhang, and Y. Liu. (2019). Combating fake news: A survey on identification and mitigation techniques. *ACM Trans. Intell. Syst. Technol.*, 10(3), 21:1–21:42.

[9] Thekumparampil, K. K., Wang, C., Oh, S., Li, L. J. (2018). Attention-based graph neural network for semi-supervised learning. *arXiv preprint*.

[10] Zhou, X., and Zafarani, R. (2019). Fake news: Fundamental theories, detection strategies and challenges. *arXiv preprint*.

Authorized licensed use limited to: Hebei University of Technology. Downloaded on January 04,2021 at 06:08:38 UTC from IEEE Xplore. Restrictions apply.