Propagation-Based Fake News Detection Using Graph Neural Networks with Transformer

Hayato Matsumoto, Soh Yoshida, and Mitsuji Muneyasu
Graduate School of Science and Engineering, Kansai University
E-mail: sohy@kansai-u.ac.jp

2021 IEEE 10th Global Conference on Consumer Electronics (GCCE) | 978-1-6654-3676-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/GCCE53005.2021.9621803

***Abstract*—The spread of fake news has become a worldwide problem, affecting public trust. Recent studies have reported that fake news and real news spread differently on social media. Thus, propagation-based detection methods, which construct graphs with users as nodes and news sharing chains as edges and simultaneously learn propagation patterns and users' preferences using Graph Neural Networks (GNNs), have attracted much attention. However, for extracting users' preferences from the graph, it is a challenge to learn the relationship between unconnected nodes. In this paper, we propose a method for fake news detection using Graph Transformer Network (GTN), which can learn efficient node representations while identifying useful connections between nodes in the original graph. The effectiveness of the proposed method is confirmed by comparison experiments using the real-world dataset composed of Twitter data.**

***Index Terms*—fake news detection, graph neural networks, propagation network, graph transformer network**

I. Introduction

Social media have become one of the primary sources of information for people around the world. Consuming news through social media is very easy and convenient. However, at the same time, there is a risk of being exposed to fake news that contains unchecked or intentionally false information aimed at misleading or manipulating readers.

Content-based and social context-based methods have traditionally been researched to detect fake news [1]. Recent studies have reported differences in the propagation patterns between fake news and real news on social platforms [2]. Specifically, fake news propagates more quickly and to a broader range of users. Based on this knowledge, propagation-based methods using Graph Neural Networks (GNNs) [3] have attracted much attention for analyzing the propagation patterns of each piece of news and the users' preferences in the sharing chain [4], [5]. However, to extract user features such as preferences from the propagation graph, it is necessary to consider the relationships between unconnected users on the graph.

In this paper, we propose a propagation-based fake news detection method using Graph Transformer Network (GTN) [6], which introduces multi-head attention and message passing mechanisms to search graphs for useful connectivity relations for identification. The contributions of this paper are summarized as follows: 1) we define a weight function and construct a weighted propagation graph to enhance the differences in propagation patterns between fake news and real news; 2) we improve the accuracy of fake news detection by introducing the GTN-based graph classification model; 3) we conduct experiments on the Twitter dataset to quantitatively confirm the effectiveness of our method.

II. Propagation Graphs Construction

We construct the propagation graph based on engagements on Twitter. Thus, the graph consists of three types of nodes: source news, tweets, and retweets. After a piece of news is published, users tweet about it. Then, users who read the tweets retweet them, and a chain of news sharing is generated. We construct an undirected tree graph $G = (V, E)$, where V denotes the nodes in the graph with $|V| = n$ and E denotes edges. The nodes are described by the feature matrix $X \in \mathbb{R}^{n \times d}$, which is the set of vectors with d dimensions. The edges are connected when news sharing occurs between the two nodes.

Fake news and real news spread at different speeds [2]. In this paper, we newly define a weight function on the edges of the graph to take this phenomenon into account. Let ΔT_{ij} denote the time difference [sec.] between the published date of node i and that of node j. Each time is converted to Unix time. In the graph, if two nodes i and j are connected, the weight a_{ij} on the edge between them is calculated as

$$a_{ij} = 1 - \frac{\Delta T_{ij}}{\max(\Delta T)}, \tag{1}$$

where $\max(\Delta T)$ is the max of the time differences. The adjacency matrix $A = [a_{ij}] \in \mathbb{R}^{n \times n}$ is used to describe the graph G.

III. Proposed Method

We aim to construct a fake news detection method using GNNs that simultaneously considers the propagation pattern of news on social media and the users' preferences. First, we generate a labeled graph dataset $\mathcal{D} = \{(G_1, y_1), (G_2, y_2), \ldots, (G_i, y_i), \ldots\}$, where $y_i \in \{0 : \text{real}, 1 : \text{fake}\}$, based on the posts crawled from social media for each news i. Second, we calculate feature matrix X combining user features and textual features that reflect the endogenous preference of each user from the past postings. Third, we obtain new node features by training GTN using the features calculated in the second step as initial node features. Finally, we aggregate the node features using the global average pooling layer and classify them using the multi-layer perceptron (MLP). The details are shown below.

A. Data Generation

In this paper, we use the FakeNewsNet dataset [7], which consists of Twitter data, to construct a labeled graph dataset $\mathcal{D}$. In [7], Shu et al. create labeled news based on two fact-checking sites, politifact.com and gossipcop.com, and collect tweets containing keywords extracted from each news. This dataset is unique because it includes both tweet ids and retweet ids, allowing us to reproduce the propagation chain. However, following Twitter's policy, Twitter data is collected via the Twitter API using the corresponding ids.

B. Features Extraction

User Features. We define the following seven types of data as user features: 1) number of words in self-introduction, 2) number of words in screen name, 3) number of followers, 4) number of friends, 5) number of tweets, 6) account verification (bool), and 7) location information (bool).

This work was supported by JST, ACT-X Grant Number JPMJAX20AN, Japan

TABLE I
PERFORMANCE COMPARISONS ON THE DATASETS OF POLITIFACT (POL) AND GOSSIPCOP (GOS)

Methods		POL				GOS			
		ACC	F1	Prec.	Rec.	ACC	F1	Prec.	Rec.
Proposed	w	**0.9379**	**0.9132**	**0.9408**	0.9795	**0.9535**	**0.9064**	0.8912	0.7661
	w/o	0.9302	0.9042	0.9404	0.9692	0.9399	0.8751	0.7855	0.8297
UPFD [4]	w	0.9146	0.8817	0.9306	0.9587	0.9517	0.8879	**0.8934**	0.7356
	w/o	0.9185	0.8902	0.9338	0.9587	0.9287	0.8715	0.7901	0.811
Geometric [5]	w	0.9146	0.8689	0.9074	**0.9897**	0.8295	0.7883	0.7378	**0.8313**
	w/o	0.9032	0.8358	0.9217	0.9637	0.8068	0.7482	0.7511	0.7438

Textual Features. Textual features are extracted by encoding the users' previous 200 tweets by using text representation learning techniques such as BERT [8]. The BERT method extracts features related to contextual expressions specific to users' preferences and fake news. We extract 768-dimensional features from pre-processed source news and tweet and retweet posts using the pre-trained BERT model such as bert-base-uncased.

C. Fake News Detection

Since most GNN models repeat the propagation and update of node features between adjacent nodes, it is difficult to consider the relationship between similar nodes in non-adjacent nodes. This effect is especially noticeable in propagation graphs with tree structures. GTN solves the problem of conventional methods by introducing multi-head attention and message passing mechanisms.

Let $H^{(l)} = \{h_1^{(l)}, h_2^{(l)}, \cdots, h_n^{(l)}\}$ denote the node features of l-th layer. Note that $H^{(0)} = X$. For the c-th head attention, the source feature $h_i^{(l)}$ and distant feature $h_j^{(l)}$ are transformed into query vector $q_{c,i}^{(l)}$ and key vector $k_{c,j}^{(l)}$ as follows:

$$\begin{aligned} q_{c,i}^{(l)} &= W_{c,q}^{(l)} h_i^{(l)} + b_{c,q}^{(l)}, \\ k_{c,j}^{(l)} &= W_{c,k}^{(l)} h_j^{(l)} + b_{c,k}^{(l)}, \\ a_{c,ij} &= W_{c,a} a_{ij} + b_{c,a}, \\ \alpha_{c,ij}^{(l)} &= \frac{\langle q_{c,i}^{(l)}, k_{c,j}^{(l)} + a_{c,ij} \rangle}{\sum_{u \in \mathcal{N}(i)} \langle q_{c,i}^{(l)}, k_{c,u}^{(l)} + a_{c,iu} \rangle}, \end{aligned} \tag{2}$$

where $\langle q, k \rangle = \exp(q^T k / \sqrt{m})$, where m is the hidden size of each head. $W_{c,q}^{(l)}$, $W_{c,k}^{(l)}$, $W_{c,a}$, $b_{c,q}^{(l)}$, $b_{c,k}^{(l)}$, $b_{c,a}$ are the trainable parameters.

After getting the graph multi-head attention, message aggregation from the distance j to the source i are made as follows:

$$\begin{aligned} v_{c,j}^{(l)} &= W_{c,v}^{(l)} h_j^{(l)} + b_{c,v}^{(l)}, \\ \hat{h}^{(l+1)} &= \|_{c=1}^{C} [\sum_{j \in \mathcal{N}(i)} \alpha_{c,ij}^{(l)} (v_{c,j}^{(l)} + w_{c,ij})], \end{aligned} \tag{3}$$

where $W_{c,v}^{(l)}$, $b_{c,v}^{(l)}$ are the trainable parameters, $\|$ is the concatenation operator, and C is the number of heads. Compare with Eq. (2), multi-head attention matrix replaces the original affinity matrix as transition matrix for message passing.

Finally, we feed $\hat{H}^{(L)} = \{\hat{h}_1^{(L)}, \hat{h}_2^{(L)}, \cdots, \hat{h}_m^{(L)}\}$, where L is the number of GTN layers, obtained from Eq. (3) into the MLP with K layers of fully connected layers to construct the fake detection model.

IV. EXPERIMENTS

In this section, we conduct experiments to verify 1) the effect of the weight function considering the time difference of news propagation and 2) the effect of propagation graph learning using GTN.

Datasets. The number of news (graphs) in the FakeNewsNet dataset are Politifact (**POL**): {real:51, fake:99} and GossipCop (**GOS**): {real: 1258, fake: 1047}, respectively.

Implementation Details. Our model consists of two GTN layers and two MLP layers. All GNN models were implemented with the PyTorch-Geometric package. We used the hidden size of GTN layers (128), optimizer (Adam), and L2 regularization weight (0.001). For evaluation metrics, we use accuracy (ACC) and f1-score (F1), respectively. All the numerical results are the results of 5-fold cross-validation with under-sampling for imbalanced data.

Baselines. We compared our proposed method, denoted as **Proposed**, with two state-of-the-art propagation-based methods using GNNs. The network architectures of each method are as follows: **UPFD** [4]: the outputs of two GCN layers and the skip-connected features of the original node features are fed into two MLP layers; **Geometric** [5]: the model consists of two GCN layers [3] and two MLP layers. For a fair comparison, all methods used the same feature matrix X.

Results. The performance comparisons on the two datasets are shown in Table I. In Table I, "w" and "w/o" represent edge-weighted and unweighted graphs, respectively, as calculated in Eq. (1). First, all the methods using edge-weighted graphs outperform the results using unweighted graphs, thus confirming the effectiveness of our graph construction method. Second, ACC and F1 of the proposed method with GTN outperformed all the results, thus confirming the effectiveness of the proposed method.

V. CONCLUSIONS

In this paper, we proposed a fake news detection method using GTN trained on weighted propagation graphs. The effectiveness of the proposed method has been confirmed by comparison experiments using real-world datasets composed of Twitter data.

REFERENCES

[1] X. Zhou and R. Zafarani, "A survey of fake news: Fundamental theories, detection methods, and opportunities," *ACM Computing Surveys*, vol. 53, no. 5, 2020.

[2] Z. Zhao, J. Zhao, Y. Sano, O. Levy, H. Takayasu, M. Takayasu, D. Li, J. Wu, and S. Havlin, "Fake news propagates differently from real news even at early stages of spreading," *EPJ Data Science*, vol. 9, no. 7, 2020.

[3] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A comprehensive survey on graph neural networks," *IEEE Trans. Neural Networks and Learning Systems*, vol. 32, no. 1, pp. 4–24, 2021.

[4] Y. Dou, K. Shu, C. Xia, P.-S. Yu, and L. Sun, "User preference-aware fake news detection," in *Proc. Int. ACM SIGIR Conf. Research and Development in Information Retrieval*, 2021.

[5] F. Monti, F. Frasca, D. Eynard, D. Mannion, and M. Bronstein, "Fake news detection on social media using geometric deep learning," in *Proc. Int. Conf. Learning Representation*, 2019.

[6] Y. Shi, Z. Huang, S. Feng, H. Zhong, W. Wang, and Y. Sun, "Masked label prediction: Unified message passing model for semi-supervised classification," *arXiv preprint arXiv:2009.03509*, 2021.

[7] K. Shu, D. Mahudeswaran, S. Wang, D. Lee, and H. Liu, "FakeNewsNet: A data repository with news content, social context, and spatiotemporal information for studying fake news on social media," *Big Data*, vol. 8, no. 3, pp. 171–188, 2020.

[8] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proc. HLT-NAACL*, 2019, pp. 4171–4186.