Information Processing and Management 58 (2021) 102712

Contents lists available at ScienceDirect

Information Processing and Management

journal homepage: www.elsevier.com/locate/ipm

Temporally evolving graph neural network for fake news detection

Chenguang Song [a], Kai Shu [b], Bin Wu [a,*]

[a] Beijing Key Laboratory of Intelligence Telecommunication Software and Multimedia, Beijing University of Posts and Telecommunications, Beijing 100876, PR China

[b] Department of Computer Science, Illinois Institute of Technology, Chicago 60601, USA

ARTICLE INFO

Keywords:
Fake news detection
Temporal evolution
Diffusion networks

ABSTRACT

The proliferation of fake news on social media has the probability to bring an unfavorable impact on public opinion and social development. Many efforts have been paid to develop effective detection and intervention algorithms in recent years. Most of the existing propagation-based fake news detection methods focus on *static* networks and assume the whole information propagation network structure is accessible before performing learning algorithms. However, in real-world information diffusion networks, new nodes and edges constantly emerge. Therefore, in this paper, we introduce a novel temporal propagation-based fake news detection framework, which could fuse structure, content semantics, and temporal information. In particular, our model can model temporal evolution patterns of real-world news as the graph evolving under the setting of continuous-time dynamic diffusion networks. We conduct extensive experiments on large-scale real-world datasets and the experimental results demonstrate that our proposed model outperforms state-of-the-art fake news detection methods.

1. Introduction

In recent years, people increasingly tend to consume news from social media platforms rather than from traditional news sources (Shu, Sliva, Wang, Tang, & Liu, 2017). The online social media platforms such as Twitter,[1] Facebook,[2] and Sina Weibo[3] have increased the ease of information propagation. However, social media platforms offer the probability for the rapid spread of misinformation and disinformation by expediting the speed and scope (Sharma et al., 2019). Compared to traditional news media, the absence of effective regulatory and fact-checking measures over posts makes fake news can be created and published online for primary motives of influencing opinions and seeking tempting profits at low cost (Bondielli & Marcelloni, 2019; Kumar & Shah, 2018). As a result, the platforms become a fertile ground for the spread of fake news (Zhang & Ghorbani, 2020).

The wide propagation of fake news will confuse and manipulate public opinions, change the way people respond to facts, and even pose a serious threat to society (Guo, Ding, Yao, Liang, & Yu, 2020). Studies in social psychology have shown that humans are irrational and vulnerable in discerning between true and false news (Zhou, Zafarani, Shu, & Liu, 2019). Thus, to increase the trustworthiness of online social networks and mitigate negative impacts caused by fake news, fake news detection and proactive intervention strategies for social media, especially after 2016 U.S. presidential election, have received lots of attention from both social network platforms and academic communities (Li, Zhang, Du, Ma, & Wang, 2021; Meel & Vishwakarma, 2020). The users of Sina Weibo can report possible fake posts to a special microblog community management center, and then the posts will be

* Corresponding author.
E-mail address: wubin@bupt.edu.cn (B. Wu).

[1] https://twitter.com/.
[2] https://www.facebook.com/.
[3] https://weibo.com/.

https://doi.org/10.1016/j.ipm.2021.102712
Received 24 April 2021; Received in revised form 8 July 2021; Accepted 1 August 2021
Available online 17 August 2021
0306-4573/

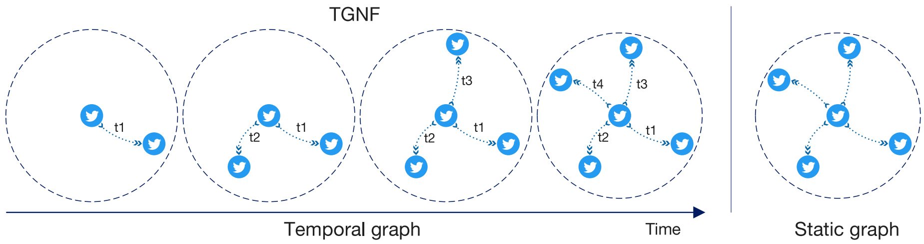


Fig. 1. Left: temporal propagation network of a piece of social media news. Each Twitter bird icon denotes a tweet. Each arrow represents a share or retweet with an associated timestamp. Right: static propagation network of a piece of social media news. Each arrow represents a share or retweet without an associated timestamp.

manually checked by professionals (Cao et al., 2018). On Facebook, users are encouraged to flag false content or news that are potentially suspicious and anomalous (Del Vicario et al., 2016). Some well-known fact-checking websites such as FactCheck.org[4] and PolitiFact.com[5] solely depend on manual identification by a small group of highly credible fact-checkers and can provide highly accurate results. However, manual fact-checking is labor-intensive and has difficulty in scaling with the volume of emerging fake news (Zhao, Da, & Yan, 2021; Zhou, & Zafarani et al., 2019). Therefore, it is important to detect fake news effectively with computational approaches.

Many existing content-based models utilize deep neural networks to learn latent textual or visual feature representation of fake news, and heavily rely on semantic information (Goldani, Momtazi, & Safabakhsh, 2021; Zhang, Fang, Qian, & Xu, 2019). However, currently content-based approaches face several challenges. First, as fake news is intentionally written to mislead readers by mimicking true news, it is difficult to efficiently detect and contrast them solely from news content (Shu, Mahudeswaran, Wang, & Liu, 2020). Second, the interpretation of the news usually lacks the necessary background knowledge such as social context, common sense (Monti, Frasca, Eynard, Mannion, & Bronstein, 2019). The currently most advanced natural language processing algorithms still fail to capture that. Previous studies have observed that fake news presents different propagation patterns from true news, which means that the propagation network of news on social networks can be leveraged to verify the given news (Shu, Mahudeswaran, Wang, Lee, & Liu, 2020; Si et al., 2020; Vosoughi, Roy, & Aral, 2018). Moreover, it is difficult for the individual users to control the spread patterns of news on social networks, which implies that propagation-based approaches may have better robustness (Monti et al., 2019). Hence, an increasing number of researchers have begun to investigated that the network of tweets and retweets relationships for each news article on social media and how it can help to infer which articles are fake in the past few years (Lu & Li, 2020; Tu et al., 2021; Vu & Jung, 2021; Wei, Xu, & Mao, 2019; Wu, Pi, Chen, Xie, & Cao, 2020). They have gained great success in the task of online fake news detection.

Despite some early promising results, most of the propagation-based fake news detection methods focus on static networks and assume the whole underlying information propagation network structure is accessible before performing learning algorithms (Jian, Li, & Liu, 2018). However, in real-world information diffusion networks, new nodes and edges constantly emerge over time. Fig. 1 illustrates the difference between *temporal* and *static* news propagation networks. As shown in Fig. 1 (left), we see that news dissemination graph is evolving in the temporal graph over time, and users spreading behaviors happening at time point t1, t2, t3, and t4. But, in Fig. 1 (right), a static graph only captures the graph structure without continuous temporal dynamic process. Recently studies have shown that temporal engagement features of users can boost the accuracy of fake news detection model (Lukasik, Cohn, & Bontcheva, 2015; Ma, Gao, Wei, Lu, & Wong, 2015; Nguyen, Sugiyama, Nakov, & Kan, 2020; Ruchansky, Seo, & Liu, 2017; Zhang, Cook, & Yilmaz, 2021; Zhou, Shu, Li, & Lau, 2019). Fig. 2 displays the average number of and cumulative average number of tweets on three real-world datasets used in this paper. Obviously, fake news and real news show differences in temporal propagation patterns. However, if applying existing fake news detection methods directly to the temporal news propagation graphs, we have to treat them as static networks by neglecting dynamic and continuously evolving nature of real-life news propagation networks (Kleinberg, 2006; Ma, Guo, Ren, Tang, & Yin, 2020; Tang, 2012). Therefore, it is necessary to develop time-aware models that help to capture the missing temporal information in static networks, and may provide unique opportunities to understand how to discriminate between fake and real news.

In this paper, a novel temporal propagation-based fake news detection framework, **T**emporally Evolving **G**raph Neural **N**etwork for **F**ake News Detection (TGNF), is proposed. First, to model temporal propagation patterns of news, we leverage the temporal graph attention neural networks (TGAT) (Rossi et al., 2020; Xu, Ruan, Korpeoglu, Kumar, & Achan, 2020) to capture its dynamic structure, content semantics, and temporal information in the process of news dissemination. Specifically, we model news propagation using the continuous-time dynamic graphs (CTDG) rather than discrete-time dynamic graphs (DTDG) (i.e., static graph snapshots), the reason of which is that the news propagation graph in social networks are naturally continuous and evolving over time as new nodes and edges are introduced to the graph continuously (Chow, Ye, Zha, & Zhou, 2018; Gomez-Rodriguez & Schölkopf, 2012; Saito, Kimura, Ohara, & Motoda, 2009; Xu, & Ruan et al., 2020). Second, inspired by adversarial learning (Goodfellow et al., 2014),

[4] http://factcheck.org/.
[5] https://www.politifact.com/.

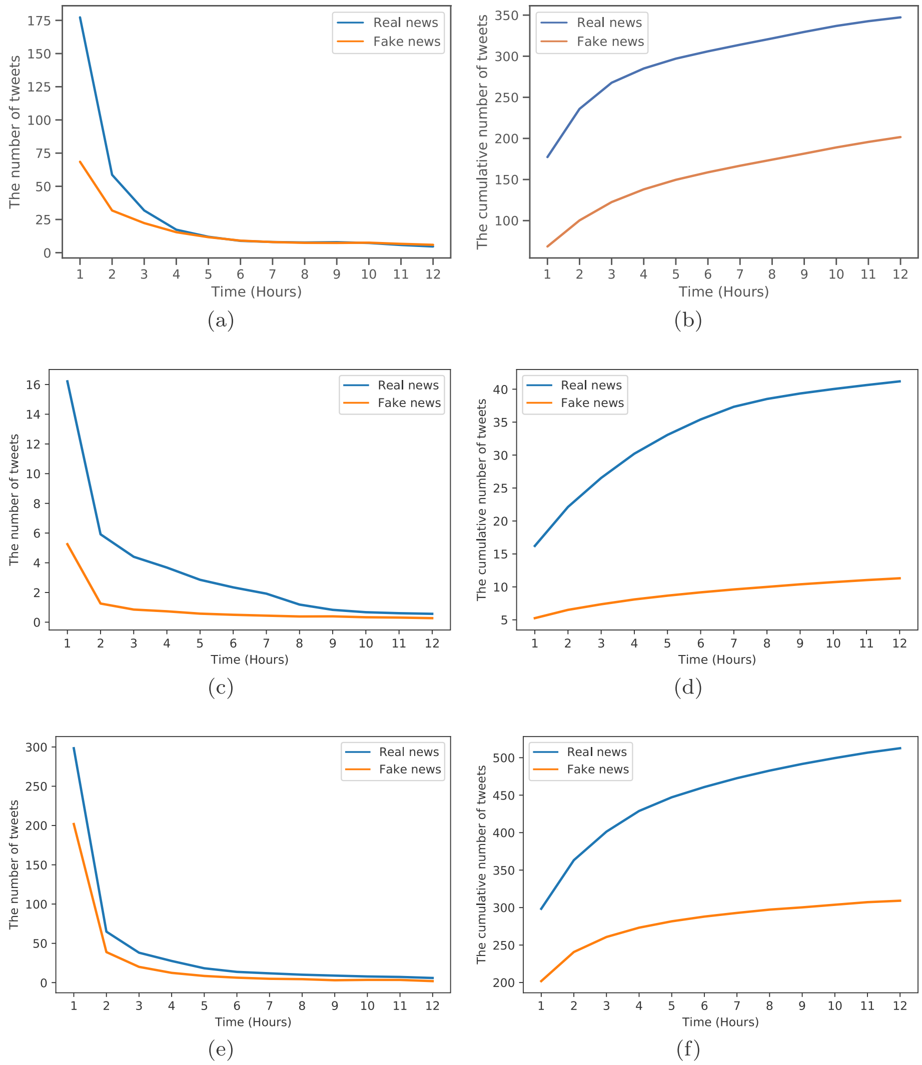


Fig. 2. (a) The average number of tweets for Weibo dataset at different timestamps; (b) The cumulative average number of tweets for Weibo dataset at different timestamps; (c) The average number of tweets for FakeNewsNet dataset at different timestamps; (d) The cumulative average number of tweets for FakeNewsNet dataset at different timestamps; (e) The average number of tweets for Twitter dataset at different timestamps; (f) The cumulative average number of tweets for Twitter dataset at different timestamps.

we designed the temporal difference network (TDN) to enable the model to concentrate on and capture the variational information between interactions rather than similar. In summary, the contributions of this paper include the following several aspects:

- In this paper, we study a novel problem of temporal propagation-based fake news detection task, which aims to fuse topological structure, content semantics, and temporal information from the perspective of continuous time.
- We propose a replacednovelnew fake news detection framework using temporal graph neural network with adversarial learning to jointly capture the temporal evolution patterns of the news diffusion, and the variational information between interactions.
- We conduct extensive experiments on three real-world datasets and experimental results show that the proposed model outperforms state-of-the-art methods.

2. Related work

2.1. Fake news detection

Though fake news is not a new phenomenon, it has been attracting increasingly public attention (Allcott & Gentzkow, 2017). The literature on fake news detection is extensive. In this section, we briefly review the existing work from the following categories: content-based and network-based and fake news detection.

Content-based Fake News Detection. For a news event, its tweets generally include a piece of text to describe it, and several attached images or videos sometimes (Guo et al., 2020). News content-based features are the most explicit clues for fake news detection, given that the social media news evaluated are primarily textual in nature (Agichtein, Castillo, Donato, Gionis, & Mishne, 2008; Bondielli & Marcelloni, 2019; Song et al., 2019). The prerequisite of news content-based fake news detection is that the content of fake news should be somewhat different from truth in some quantifiable way (Sharma et al., 2019; Zhou, & Zafarani et al., 2019). Some early studies investigated the linguistic features of news content such as lexical features, syntactic features, and topic features, then designed a set of manual linguistic cues to determine the authenticity of the given news (Castillo, Mendoza, & Poblete, 2013; Rubin, Chen, & Conroy, 2015; Sejeong, Meeyoung, Kyomin, Wei, & Yajun, 2013). However, these methods have difficulty not only in generalizing hand-crafted linguistic features across topics, languages, and domains but also in utilizing the rich semantic and contextual information (Sharma et al., 2019).

To address the drawbacks of linguistics-based methods, the deep neural networks-based methods such as recurrent neural network (RNN) (Alkhodair, Ding, Fung, & Liu, 2020; Chen, Li, Yin, & Zhang, 2018; Liu & Wu, 2018), convolutional neural networks (CNN) (Ajao, Bhowmik, & Zargari, 2018; Goldani, Safabakhsh, & Momtazi, 2021; Yu, Liu, Wu, Wang, & Tan, 2017), variational autoencoders (VAEs) (Cheng, Nazarian, & Bogdan, 2020; Khattar, Goud, Gupta, & Varma, 2019), and attention mechanism (Chen, Sui, Hu, & Gong, 2019; Guo, Cao, Zhang, Guo, & Li, 2018) have been widely explored in recent years, because these methods can automatically learn latent textual representation and capture complex contextual patterns of news content. Ma et al. presented the first work that suggests the use of deep learning techniques for identifying fake news (Ma et al., 2016). The input of their method is tf-idf feature but it shows better performance than the methods leveraging hand-crafted features. Wang et al. proposed a CNN-based model to classify fake news (Wang, 2017). Specifically, it obtains the word vector from pre-trained word2vec embeddings (Mikolov, Sutskever, Chen, Corrado, & Dean, 2013) and experimental results show that the method achieves better detection accuracy than traditional machine learning-based methods. Several studies attempt to leverage the conflicting viewpoints to facilitate the detection (Tian et al., 2020; Zubiaga et al., 2018). Inspired by adversarial learning (Goodfellow et al., 2014), Ma et al. proposed a generative adversarial networks (GAN)-style method, which use an additional text generator to pressurize the discriminator to learn strong indicative representations (Ma, Gao, & Wong, 2019). By exploiting the user responses to a target claim, stance information is proved to be a strong indicator for fake news detection (Dungs, Aker, Fuhr, & Bontcheva, 2018; Kochkina, Liakata, & Zubiaga, 2018; Ma, Gao, & Wong, 2018a). Recently, scholars also explored news content-based fake news detection in various way such as domain adaption (Silva, Luo, Karunasekera, & Leckie, 2021), knowledge enhance (Cui et al., 2020; Dun, Tu, Chen, Hou, & Yuan, 2021), temporal pattern analysis (Ruchansky et al., 2017), and weak supervision learning (Liu & Wu, 2020; Shu, & Zheng et al., 2020).

A group of recent approaches utilizes visual cues extracted from the attached images or videos (Cui, Wang, & Lee, 2019; Vishwakarma, Varshney, & Yadav, 2019; Xu, Zeng, & Mao, 2020). News with visual information is likely to attract much more attention from social media users and thus gains a greater range of information dissemination (Qi, Cao, Yang, Guo, & Li, 2019). Jin et al. first proposed a RNN-based automatic multimodal fake news detection model to fuse the visual and textual information of the post using an attention mechanism (Jin, Cao, Guo, Zhang, & Luo, 2017). Wang et al. proposed a multi-task learning model to learn textual and visual transferable feature representations among all the posts with the goal of doing with non-transferable event-specific features (Wang et al., 2018). Zhou et al. presented a novel fake news method considering the correlations across the modalities (Zhou, Wu, & Zafarani, 2020).

Network-based Fake News Detection. Network-based fake news detection utilizes lots of interactions among users or content such as commenting, retweeting, and following in news propagation (Huang, Zhou, Wu, Wang, & Wang, 2019; Liu, Jin, & Shen, 2019; Yuan, Ma, Zhou, Han, & Hu, 2020). Understanding the propagation patterns of fake news is of paramount importance as it provides useful insights for identification of fake news (Silva, Han, Luo, Karunasekera, & Leckie, 2020; Zannettou, Sirivianos, Blackburn, & Kourtellis, 2019; Zhou, Xiu, Wang, & Yu, 2021). Both homogeneous networks and heterogeneous networks can be constructed to model the propagation of news (Shu et al., 2017).

Homogeneous networks consist of a single type of nodes and edges (Zhou & Zafarani, 2019). By analyzing the diffusion of false and true news in Twitter in eleven years from the perspective of homogeneous graph, Vosoughi et al. find that false news propagates faster, farther, and more broadly when compared to truth (Vosoughi et al., 2018). Ma et al. adopted top-down and bottom-up tree-structured RNN to incorporate textual features and propagation structure features (Ma, Gao, & Wong, 2018b). Similarly, Bian et al. proposed a novel bi-directional graph convolutional network model to learn the representation of content semantics and diffusion graph of news (Bian et al., 2020).

Heterogeneous networks contain multiple types of nodes or edges (Zhou & Zafarani, 2019). Yuan et al. jointly encodes the local semantic and global structure of the diffusion graph based on a heterogeneous graph constructed posts, comments, and users (Yuan, Ma, Zhou, Han, & Hu, 2019). Huang et al. proposed a meta-path-based heterogeneous graph attention network framework to capture global semantic relations of text contents (Huang, Yu, Wu, & Wang, 2020). To improve robustness of graph-based detector, Yang et al. first model the rich information of entities through a heterogeneous information network, and then use a special graph adversarial learning framework to force their model to learn more distinctive structure features (Yang et al., 2020). Nguyen et al. proposed an inductive heterogeneous graph representation framework, Factual News Graph (FANG), which can effectively exploit social structure and engagement patterns of users for fake news detection (Nguyen et al., 2020).

In general, our research falls into the homogeneous network-based fake news detection. Different from all the aforementioned work based on *static* networks, our study aims to classify social media news from the perspective of temporal diffusion networks (i.e., *dynamic* networks).

2.2. Dynamic Graph Neural Networks

In recent years, we have witnessed many successful deep graph learning techniques (i.e., Graph Neural Networks (GNN)) as their ability to model complex relationships and inter-dependencies on graphs (Han et al., 2021). The nodes in the network represent entities, and the edges indicate relationships among those entities (Holme, 2015). Many real-life complex systems, such as social networks, recommender systems, can be characterized by complex dynamic networks (Gergely, Albert-László, & Tamás, 2007; Holme & Saramäki, 2012). However, most of the prevalent GNN-based models have assumed that the underlying graph is static while ignored its temporal evolution (Rossi et al., 2020). Hence, it would be of special advantages to perform inference on graphs in a temporal dynamic manner.

The dynamic graphs generally include discrete-time dynamic graphs (DTDG) and continuous-time dynamic graphs (CTDG) (Kazemi et al., 2020). In particular, DTDG are usually represented as a sequence of static graph snapshots at different time steps. The basic idea of the DTDG-based deep graph representation learning algorithms is to learn node embedding by aggregating the information of graph snapshots (Lu, Wang, Shi, Yu, & Ye, 2019; Manessi, Rozza, & Manzo, 2020; Sankar, Wu, Gou, Zhang, & Yang, 2020). The CTDG-based approaches aim at capturing the temporal evolution pattern of the network and dynamically learning node embedding in continuous time (Kumar, Zhang, & Leskovec, 2019; Trivedi, Farajtabar, Biswal, & Zha, 2019; Zhang et al., 2020).

Most of the existing CTDG-based neural network models are more suitable for the task of temporal node classification on a single dynamic graph. However, the propagation-based fake news detection task is usually formulated as the task of temporal graph classification on different dynamic graphs. Considering the differences between them, we proposed our method by introducing a special TDN and graph convolutional network to improve the existing continuous graph neural network models.

3. Problem formulation

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be an unweighted static graph representing a news propagation network. $\mathcal{V} = \{v_1, v_2, \dots, v_{\mathcal{N}}\}$ is the set of nodes and $\mathcal{E} = \{e_1, e_2, \dots, e_{\mathcal{M}}\}$ is the set of edges, where v_i represents a tweet, $\mathcal{N}$ represents the number of relevant tweets in $\mathcal{G}$, and $\mathcal{M}$ denotes the number of the observed interaction events. $\mathbf{h}_i \in \mathbb{R}^d$ is the feature representation of tweet v_i. Each edge $e_{ij} \in \mathcal{E}$ denotes node v_i has a response to v_j, and also can be formulated as an unweighted adjacency matrix $\mathcal{A} = (a_{ij})_{\mathcal{N}\times\mathcal{N}}$, where

$$a_{ij} = \begin{cases} 1 & if\ e_{ij} \in \mathcal{E} \\ 0 & otherwise \end{cases} \tag{1}$$

Similar to static graph, a continuous dynamic news propagation network $\mathcal{G}(t) = (\mathcal{V}(t), \mathcal{E}(t))$ consists of nodes set $\mathcal{V}(t) = \{v_1^{(t_1)}, v_2^{(t_2)}, \dots, v_{\mathcal{N}(t)}^{(t_{\mathcal{N}(t)})}\}$ and edges set $\mathcal{E}(t) = \{e_1^{(t_1)}, e_2^{(t_2)}, \dots, e_{\mathcal{M}(t)}^{(t_{\mathcal{M}(t)})}\}$ at time t. Each node $v_i^{(t_i)} \in \mathcal{E}(t)$ indicates that the tweet v_i is published as time t_i. Each edge $e_{ij}^{(t_x)} \in \mathcal{E}$ means that node v_i has a response to v_j at time point t_x. Specifically, to learn the temporal representation of each node, the response behavior (i.e., $e_{ij}^{(t_x)}$) is modeled as an interaction event between node v_i and node v_j. $\mathcal{N}(t) = |\mathcal{V}(t)|$ is the total number of tweets at time t. $\mathcal{M}(t) = |\mathcal{E}(t)|$ is the total number of interaction events (i.e., reply or retweet) at time point t. $\mathbf{h}_i(t) \in \mathbb{R}^d$ is the feature representation of tweet v_i at time t. $\mathcal{A}(t) = (a_{ij}(t))_{\mathcal{N}(t)\times\mathcal{N}(t)}$ is the adjacency matrix of $\mathcal{G}(t)$ at time point t, where

$$a_{ij}(t) = \begin{cases} 1 & if\ e_{ij} \in \mathcal{E}(t) \\ 0 & otherwise \end{cases} \tag{2}$$

When $t = t_{max}$, $\mathcal{G} = \mathcal{G}(t_{max})$, $\mathcal{V} = \mathcal{V}(t_{max})$, $\mathcal{A} = \mathcal{A}(t_{max})$, and $\mathcal{N} = \mathcal{N}(t_{max})$. $\mathcal{G}(t)$ is associated with a ground-truth label $y \in \{0, 1\}$ describing its veracity, where $y = 0$ indicates $G(t)$ is true news, and $y = 1$ means $G(t)$ is fake news. We formulate the fake news detection problem in this paper as follows.

Problem Definition: Given a temporal news propagation graph $\mathcal{G}(t) = (\mathcal{V}(t), \mathcal{E}(t))$, the goal is to learn a mapping function $\mathcal{F} : \mathcal{F}(\mathcal{G}(t)) \rightarrow \hat{y}$ to classify the veracity labels of $\mathcal{G}(t)$ by tracking the corresponding chronological interaction events.

4. Model

4.1. Model framework

In this subsection, we provide a brief model framework help readers to understand the proposed method easily. Fig. 3 shows how the model works in three different time points (i.e., t_1, t_2, t_3). The models at different time points are connected in series, for example, the model at time point t_2 is obtained by continuing to train on the basis of the model at time point t_1. The main goal of this study is to build a temporal or streaming fake news detection model. More exactly, we aim to model dynamic propagation patterns of social media news along with new nodes join the network continuously and new edges being created at any time. Therefore, the model should be able to output the prediction results at any time point. For each time point, the model first produces the raw feature representation (i.e., $\mathcal{H}(t) = [\mathbf{h}_i(t), \dots, \mathbf{h}_{\mathcal{N}(t)}(t)]$) of each node of $\mathcal{G}(t)$ using input embeddings. Then, the sum of $\mathcal{H}(t)$ and $\mathcal{S}(t) = \{\mathbf{s}_1(t), \dots, \mathbf{s}_{\mathcal{N}(t)}(t)\}$ is fed into the temporal graph attention network to obtain the nodes' temporal feature representation of $\mathcal{G}(t)$ at time point t (i.e., $\tilde{\mathcal{H}}(t) = [\tilde{\mathbf{h}}_i(t), \dots, \tilde{\mathbf{h}}_{\mathcal{N}(t)}(t)]$), where $\mathcal{S}(t)$ represents the memory vectors of nodes, and $\bigoplus$ means additive operation. The temporal memory module store temporal interaction information between different nodes and update the memory vectors of these nodes. Next, $\tilde{\mathcal{H}}(t)$ is piped to the graph convolutional layer (GCL) to update the information of the neighbor nodes. At last, we average the embeddings of the nodes with the mean pooling operation, followed by a feed-forward neural network (FFN) layer and a softmax layer for prediction. Specifically, we design the TDN to help temporal graph attention network to focus on the variational information between interactions. $\mathcal{L}_d$ is the loss of the TDN.

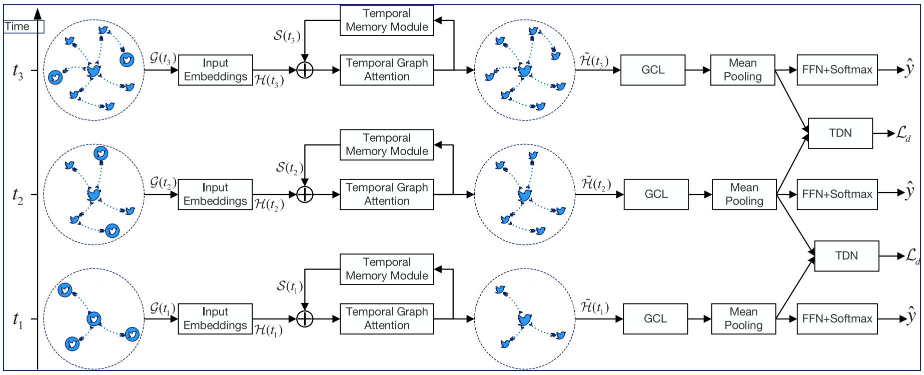


Fig. 3. The proposed model framework.

4.2. Temporal embedding network

Recently, there has been growing interest in generalizing representation learning techniques to temporal graphs. Among them, TGAT is an effective general inductive representation learning framework for dynamic graphs. TGAT is first proposed by Xu, and Ruan et al. (2020), and then further improved by Rossi et al. (2020). In this subsection, we describe the single-layer TGAT network architecture without considering edge features. In particular, to improve the training efficiency, TGAT is trained with batches of interaction data rather than learn node representation from a sequence of interactions by processing one interaction after the other, which follows previous work (Kumar et al., 2019). If you are interested in the full model, especially batch training, we refer the reader to Kumar et al. (2019), Rossi et al. (2020) and Xu, and Ruan et al. (2020) for a more detailed explanation of the model. *The input and output of the Temporal Embedding Network are a temporal news diffusion graph* $\mathcal{G}(t)$ and corresponding time-aware representation $\tilde{\mathcal{H}}_i(t)$, respectively.

Input Embeddings. The input embedding layer aims to produce a raw feature representation for a given tweet (i.e., node $v_i^{(t_i)}$). First, we convert each word of $v_i^{(t_i)}$ to a sequence of pretrained word vectors:

$$[\boldsymbol{\omega}_1, \dots, \omega_j, \dots] \leftarrow \mathrm{WordEmbed}(v_i^{(t_i)}) \tag{3}$$

where $\omega_j \in \mathbb{R}^d$. Second, we average the embeddings of the word vectors to obtain the raw node feature representation $\mathbf{h}_i(t) \in \mathbb{R}^d$ of node $v_i^{(t_i)}$.

Time Encoding Function. The functional time encoding is an important theoretical basis of temporal graph attention. Inspired by the classic harmonic analysis, Xu et al. convert the challenge of learning functional time encoding to the kernel and distributional learning problems (Xu, & Ruan et al., 2020). It can map time from the time domain to the d_t dimensional vector space (Kazemi et al., 2019; Li et al., 2020; Rossi et al., 2020), and learn representations on temporal graphs by combining with self-attention mechanism (Vaswani et al., 2017). Same to (Rossi et al., 2020; Xu, & Ruan et al., 2020), the time encoding function of given time t can be defined as follows.

$$\boldsymbol{\varphi}(t) = \cos\left(\mathbf{w}_t \times t + \mathbf{b}_t\right) \tag{4}$$

where $\mathbf{w}_t \in \mathbb{R}^{d_t}$ and bias term $\mathbf{b}_t$ are learned parameters, and $\boldsymbol{\varphi}(t) \in \mathbb{R}^{d_t}$.

Temporal Graph Attention. We provide a brief introduction to TGAT used in this paper. Similar to GraphSAGE (Hamilton, Ying, & Leskovec, 2017) and GAT (Veličković et al., 2018), the TGAT layer can be considered to be a local aggregation operator. The input of TGAT layer is features and timestamps of a node and its temporal neighborhood, and its output is the time-aware representation for the node at time t. Given a target node $v_i^{(t_i)}$, its temporal neighbor nodes is defined as follows.

$$\mathrm{Neigh}\left(v_i^{(t_i)}; t\right) = \left\{v_{i1}^{(t_{i1})}, v_{i2}^{(t_{i2})}, \dots, v_{ib}^{(t_{ib})}\right\} \tag{5}$$

where, b is the number of its temporal neighbor nodes. Here, $\mathbf{h}_i(t) \in \mathbb{R}^d$ is the features of node $v_i^{(t_i)}$ at time t, and $\left\{\mathbf{h}_{i1}\left(t_1\right), \mathbf{h}_{i2}\left(t_2\right), \dots, \mathbf{h}_{ib}\left(t_b\right)\right\} \in \mathbb{R}^d$ $(t_N \le t)$ is the features of its neighbor nodes at time t. As the TGAT only pays attention to the timespan (i.e., relative time difference), according to Eq. (4), the time encodings of node $v_i^{(t_i)}$ and its neighbor nodes are denoted by $\boldsymbol{\varphi}(0) \in \mathbb{R}^{d_t}$, and $\left\{\varphi\left(t-t_1\right), \varphi\left(t-t_2\right), \dots, \varphi\left(t-t_b\right)\right\} \in \mathbb{R}^{d_t}$, respectively. Different from the graph attention network (GAT) proposed by Veličković et al. (2018), TGAT adopt the scaled dot-product attention. The Queries, Keys and Values can be defined as follows.

$$\begin{cases} \mathbf{Q}(t) = [\mathbf{h}_i(t) \parallel \boldsymbol{\varphi}(0)]\mathbf{W}_Q \\ \mathbf{K}(t) = [\mathbf{h}_{i1}\left(t_1\right) \parallel \varphi\left(t-t_1\right), \mathbf{h}_{i2}\left(t_2\right) \parallel \varphi\left(t-t_2\right), \dots, \mathbf{h}_{ib}\left(t_b\right) \parallel \varphi\left(t-t_b\right)]\mathbf{W}_K \\ \mathbf{V}(t) = [\mathbf{h}_{i1}\left(t_1\right) \parallel \varphi\left(t-t_1\right), \mathbf{h}_{i2}\left(t_2\right) \parallel \varphi\left(t-t_2\right), \dots, \mathbf{h}_{ib}\left(t_b\right) \parallel \varphi\left(t-t_b\right)]\mathbf{W}_V \end{cases} \tag{6}$$

where $\|$ is concatenate operation, $\mathbf{W}_Q \in \mathbb{R}^{(d_t+d)\times d_h}$, $\mathbf{W}_K \in \mathbb{R}^{(d_t+d)\times d_h}$ and $\mathbf{W}_V \in \mathbb{R}^{(d_t+d)\times d_h}$ are trainable weight matrices, which are designed to capture the interactions between time encoding and node features. According to self-attention mechanism proposed by Vaswani et al. (2017), the self-attention function of TGAT is defined as follows.

$$\mathrm{Attention}(\mathbf{Q}(t),\mathbf{K}(t),\mathbf{V}(t)) = \mathrm{softmax}(\mathbf{Q}(t)\times\mathbf{K}(t)^{\top}/\sqrt{d_t+d})\times\mathbf{V}(t) \tag{7}$$

The output of above self-attention function is the hidden neighborhood representations $\hat{\mathbf{h}}_i(t) \in \mathbb{R}^{d_h}$. To combine the aggregated information with the target node representation, the target node features and the neighborhood representation are fed into an FFN layer to obtain the final time-aware representation $\tilde{\mathbf{h}}_i(t)$ of the target node $v_i^{(t_i)}$ at time t:

$$\begin{aligned}\tilde{\mathbf{h}}_i(t) &= \mathrm{FFN}(\hat{\mathbf{h}}_i(t) \,\|\, (\mathbf{h}_i(t) \,\|\, \boldsymbol{\varphi}(0)))\\ &= \mathrm{ReLU}([\hat{\mathbf{h}}_i(t) \,\|\, (\mathbf{h}_i(t) \,\|\, \boldsymbol{\varphi}(0))]\mathbf{W}_i + \mathbf{b}_i)\acute{\mathbf{W}}_i + \acute{\mathbf{b}}_i\end{aligned} \tag{8}$$

where $\mathbf{W}_i \in \mathbb{R}^{(d_t+d+d_h)\times d_f}$, $\acute{\mathbf{W}}_i \in \mathbb{R}^{d_f\times d_h}$, $\mathbf{b}_i$ and $\acute{\mathbf{b}}_i$ are bias term, and $\tilde{\mathbf{h}}_i(t) \in \mathbb{R}^{d_h}$. Note that $d_f = d_t = d_h = d$ in this paper and Eq. (7) can extended to the multi-head setting. Hence, the self-attention function of jth head is:

$$\hat{\mathbf{h}}_i^{(j)}(t) = \mathrm{Attention}^{(j)}(\mathbf{Q}(t),\mathbf{K}(t),\mathbf{V}(t)) \qquad j = 1,2,\ldots,k \tag{9}$$

Then, under the setting of the multi-head attention, Eq. (8) is also changed to:

$$\begin{aligned}\tilde{\mathbf{h}}_i(t) &= \mathrm{FFN}(\hat{\mathbf{h}}_i(t) \,\|\, (\mathbf{h}_i(t) \,\|\, \boldsymbol{\varphi}(0)))\\ &= \mathrm{FFN}[(\hat{\mathbf{h}}_i^{(1)}(t) \,\|\, \cdots \,\|\, \hat{\mathbf{h}}_i^{(k)}(t)) \,\|\, (\mathbf{h}_i(t) \,\|\, \boldsymbol{\varphi}(0))]\end{aligned} \tag{10}$$

Temporal Memory Module. We introduce the details of temporal memory module (TMM) proposed by Rossi et al. (2020). TMM assumes that, for each node $v_i^{(t_i)}$, there exists a memory vector $\mathbf{s}_i(t)$ to store history interactive memory in a compressed format. It should be noted that memory of each node is initialized as a zero vector. The memory vector $\mathbf{s}_i(t)$ is updated after node $v_i^{(t_i)}$ interact with another node. For example, an interaction event $e_{ij}(t)$ between node $v_i^{(t_i)}$ and node $v_j^{(t_j)}$ at time t can be saved as a message vector:

$$\mathbf{m}_i(t) = [\mathbf{s}_i(t^-) \,\|\, \mathbf{s}_j(t^-) \,\|\, \varphi(\Delta t)] \in \mathbb{R}^{(2d+d_t)} \tag{11}$$

where $\|$ is concatenate operation, $\mathbf{s}_i(t^-) \in \mathbb{R}^d$ and $\mathbf{s}_j(t^-) \in \mathbb{R}^d$ are the memory vector of node i and node j before time t, Δt is the timespan between node $v_i^{(t_i)}$ and node $v_j^{(t_j)}$ (i.e., $\Delta t = |t_j - t_i|$). Then, the memory vector of node $v_i^{(t_i)}$ at time t can be updated by a memory update function:

$$\mathbf{s}_i(t) = \mathrm{GRU}(\mathbf{m}_i(t), \mathbf{s}_i(t^-)) \in \mathbb{R}^d \tag{12}$$

In order to improve computational efficiency, TMM uses batch processing to capture memory vectors of b previous interaction events at the same time, which can be defined as an aggregation function:

$$\bar{\mathbf{m}}_i(t) = \mathrm{Agg}(\mathbf{m}_i(t_1),\ldots,\mathbf{m}_i(t_b)) \in \mathbb{R}^{(2d+d_t)} \tag{13}$$

where $t_1,\ldots,t_b \le t$. In particular, aggregation operation can keep only most recent message vectors or average all message vectors for a given node $v_i^{(t_i)}$. In fact, experimental results of the above two ways are similar, which is consistent with the findings of Rossi et al. (2020). Therefore, under the setting of batch processing, the memory update function is changed to:

$$\mathbf{s}_i(t) = \mathrm{GRU}(\bar{\mathbf{m}}_i(t), \mathbf{s}_i(t^-)) \in \mathbb{R}^d \tag{14}$$

The memory vectors the nodes in $\mathcal{G}(t)$ can be represented as $\mathcal{S}(t) = \{\mathbf{s}_1(t),\ldots,$ $\mathbf{s}_{\mathcal{N}(t)}(t)\}$. We have pointed out that the memory vectors of each node is initialized as a zero vector. For example, when a node $v_i^{(t_i)}$ appears for the first time, the feature of $v_i^{(t_i)}$ input to TGAT layer is its raw features (i.e., $\mathbf{h}_i(t) \in \mathbb{R}^d$), and for all the nodes of $\mathcal{G}(t)$, it can be denoted as:

$$\boldsymbol{\mathcal{H}}(t) \leftarrow \boldsymbol{\mathcal{H}}(t) + \mathcal{S}(t) \tag{15}$$

When node $v_i^{(t_i)}$ appears in next batch at time point t, the features are:

$$\tilde{\mathbf{h}}_i(t) \leftarrow \tilde{\mathbf{h}}_i(t) + \mathbf{s}_i(t) \tag{16}$$

Therefore, for all the nodes of $\mathcal{G}(t)$, it can be denoted as:

$$\tilde{\boldsymbol{\mathcal{H}}}(t) \leftarrow \tilde{\boldsymbol{\mathcal{H}}}(t) + \mathcal{S}(t) \tag{17}$$

As previously discussed, TGAT is trained with batches of interaction data, as opposed to individual interaction. For each batch, TGAT will produce or update the node embeddings of the temporal graph $\mathcal{G}(t)$, which can be denoted as:

$$\tilde{\mathcal{H}}_i(t) = (\tilde{\mathbf{h}}_1(t), \tilde{\mathbf{h}}_2(t), \ldots, \tilde{\mathbf{h}}_{\mathcal{N}(t)}(t)) \tag{18}$$

where $\tilde{\mathcal{H}}_i(t)$ represents the node embeddings of temporal graph $\mathcal{G}(t)$ after processed by the ith batch. Noticeably, the same node may involve multiple interactions in the same batch. As TGAT could process multiple interactions (i.e, b interactions (see Eq. (6)) in a

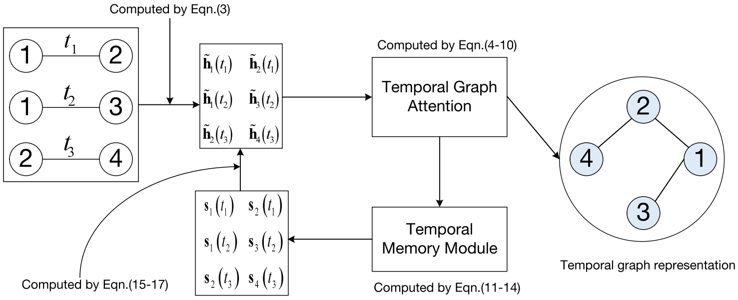


Fig. 4. The flow of operations of temporal embedding network.

batch simultaneously, we just need to choose the output of the latest one as node embeddings. In this paper, batch size equals the number of temporal neighbor nodes b. Note that the batch size of TGAT indicates the number of the interactions in a batch for a piece of news, instead of the number of news. Assuming that there exist n batches for each news, the final node embeddings of $\mathcal{G}(t_{max})$ is $\tilde{\mathcal{H}}_n(t_{max})$. Fig. 4 shows the computations performed by temporal embedding network on a bath of training examples. For each interaction at some time point t_i, the network first produces the feature representation of each node ($\tilde{\mathbf{h}}_j(t_i)$) using input embeddings (via Eq. (3)) and just updated memory (via Eq. (16)). Since node 1 and node 2 appear in the same batch twice, $\mathbf{s}_1(t_1) = \mathbf{s}_1(t_2)$ and $\mathbf{s}_2(t_1) = \mathbf{s}_2(t_3)$, and $\tilde{\mathbf{h}}_1(t_1) = \tilde{\mathbf{h}}_1(t_2)$ and $\tilde{\mathbf{h}}_2(t_1) = \tilde{\mathbf{h}}_2(t_3)$. Then, $\tilde{\mathbf{h}}_*(t_*)$ are fed into TGAT, which outputs an embedding for each node by encoding temporal interaction events. At last, each interaction between nodes is used to update the corresponding nodes' memory.

4.3. Temporal difference network

As the training of TGAT network using batches of data, intuitively, we hope TGAT to pay close attention to variational information between batches rather than similar. Inspired by adversarial learning (Goodfellow et al., 2014), we designed the TDN to force TGAT to focus on the information. First, after the temporal embedding network embed node representation of $\mathcal{G}(t)$ for batch i, we employ mean-pooling operators to aggregate information from $\tilde{\mathcal{H}}_i(t)$, which can be formulated as follows.

$$\mathbf{S}_i = \mathrm{MeanPooling}(\tilde{\mathcal{H}}_i(t)) \tag{19}$$

where MeanPooling indicates mean-pooling operation. The temporal difference loss between batch i and $i+1$ is defined as:

$$\mathcal{L}_i = \psi(\mathbf{S}_i, \mathbf{S}_{(i+1)}) \tag{20}$$

where $\psi(\cdot)$ denotes cosine similarity measure, and $1 \le i \le n-1$. Thus, the loss of the TDN can be denoted as:

$$\mathcal{L}_d = \frac{1}{n-1} \sum_{1 \le i \le n-1} \mathcal{L}_i \tag{21}$$

4.4. News predictor

We have introduced the main modules of this paper in the previous section. Next, we first average the nodes' embeddings $\tilde{\mathcal{H}}_n(t_{max})$:

$$\mathbf{S}_{max} = \mathrm{MeanPooling}(\tilde{\mathcal{H}}_n(t_{max})) \in \mathbb{R}^d \tag{22}$$

where MeanPooling indicates the mean-pooling operation. Then, $\mathbf{S}_{max}$ is fed into an FFN layer and a softmax layer to make a prediction. Therefore, the news predictor is defined as:

$$\hat{y} = \mathrm{softmax}\left(\sigma[\mathbf{S}_{max}\mathbf{W}_{max} + \mathbf{b}_{max}]\right) \tag{23}$$

where $\mathbf{W}_{max} \in \mathbb{R}^{d\times 2}$ is trainable weight matrices, $\mathbf{b}_{max}$ is a bias term, and $\sigma(\cdot)$ is the TanH activation function. $\hat{y} = [\hat{y}_0, \hat{y}_1]$ denotes the probability of given a piece of news is true (i.e., $\hat{y}_0 = 0$) or fake (i.e., $\hat{y}_1 = 1$). For each news, we adopt binary cross-entropy loss function to define the loss function $\mathcal{L}_c(\Theta_c)$ as follows.

$$\mathcal{L}_c(\Theta_c) = -y\log(\hat{y}_1) - (1-y)\log(\hat{y}_0) \tag{24}$$

where Θ_c is the learned parameters of the model.

4.5. Graph convolutional layer

In our model, a node usually represents a tweet, retweet, or reply, and the interaction between each pair of nodes only happens once. Accordingly, the newly emerging interaction could change the topological structure of the graph and influence their neighbor nodes. However, it is undesirable to update the information of the neighbor nodes for each interaction, which will largely increase the computational complexity of the model and contrary to the purpose of batch processing (Kumar et al., 2019). In order to alleviate the drawbacks, we have adopted a trade-off strategy that the information diffusion graphs of news could be fed into a two layer of GCL (see Appendix A) to update the information of the influenced nodes *when the temporal embedding network has already converged in the training procedure*. It can be formulated as:

$$\begin{cases} \tilde{\mathcal{H}}_n(t_{max}) = \mathrm{GCL}^{(1)}(\mathcal{G}(t_{max}), \tilde{\mathcal{H}}_n(t_{max})) \in \mathbb{R}^{d} \\ \tilde{\mathcal{H}}_n(t_{max}) = \mathrm{GCL}^{(2)}(\mathcal{G}(t_{max}), \tilde{\mathcal{H}}_n(t_{max})) \in \mathbb{R}^{d_g} \end{cases} \tag{25}$$

In fact, GAT is also acceptable, and their performance is very similar. The reason why we set the layer of GCL as 2 is that it shows better and stable performance. Then, the news predictor can be defined as:

$$\hat{y} = \log\mathrm{softmax}\left(\tilde{\mathcal{H}}_n(t_{max})\mathbf{W}_{gcl} + \mathbf{b}_{gcl}\right) \tag{26}$$

where $\mathbf{W}_{gcl} \in \mathbb{R}^{d_g\times 2}$ is the learned transformation matrix, and $\mathbf{b}_{gcl}$ is bias vector. Its binary cross-entropy loss function is defined as follows.

$$\mathcal{L}_g\left(\Theta_g\right) = -y\log\left(\hat{y}_1\right) - (1-y)\log\left(\hat{y}_0\right) \tag{27}$$

where Θ_g is the parameters of the GCL layers, and $\hat{y} = \left[\hat{y}_0, \hat{y}_1\right]$ denotes the probability of given a piece of news is true (i.e., $\hat{y}_0 = 0$) or fake (i.e., $\hat{y}_1 = 1$).

4.6. Model integration

In this work, we adopt two-stage training strategy to learn the proposed model's parameters. The first step is to minimize the following loss function:

$$\mathcal{L}(\Theta_c) = \mathcal{L}_c(\Theta_c) + \lambda \times \mathcal{L}_d \tag{28}$$

where λ is a hyper-parameter and used to balance the importance between $\mathcal{L}_c$ and $\mathcal{L}_d$. *When the temporal embedding network has already converged, we freeze the parameters of temporal embedding network and remove TDN, and then feed the output of temporal embedding network to GCL layer*. Next, by minimizing the loss of Eq. (27), we can get the final prediction results for each news. In this paper, we adopt the stochastic gradient algorithm and choose Adam as the optimizer to train and to optimize the proposed framework.

5. Experiments

In the following subsections, we first provide a brief introduction of the datasets used in the experiments. Second, we describe the model settings. Third, we introduce a series of state-of-the-art baseline fake news detection approaches. Finally, we make comparisons between the model and baseline methods on three datasets, and then bring an detail analysis for experimental results.

5.1. Datasets

We conduct extensive experiments on three representative real-world and publicly available datasets constructed from Twitter and Sina Weibo. These datasets contain the temporal information, propagation path, and text content. Some important statistics of three datasets are summarized in Table 1.

- Weibo: This dataset is first presented in Ma et al. (2016) for rumor classification, and crawled from Sina Weibo, the most popular social media site in China. The raw dataset consists of 2351 ture news and 2312 fake news. Due to limited GPU resource, we removed some news with nodes more than 2000. After removing these news, the number of news actually used in the experiments is shown in Table 1.
- FakeNewsNet: The dataset is developed by Shu, and Mahudeswaran et al. (2020). The news content is crawled from two fact-checking platforms: GossipCop[6] and PolitiFact.[7] The tweets related to a news are collected from Twitter API. We removed the news with missing text and timestamp. After preprocessing the dataset, the number of news actually used in this work is shown in Table 1.
- Twitter: The Twitter dataset is released by Ma, Gao, and Wong (2017). In fact, it includes two datasets (i.e., twitter15 and twitter 16). We selected non-rumors and true rumors of twitter15 and twitter16 as real news and fake news, respectively. We preprocessed Twitter dataset with the way same to Weibo and FakeNewsNet datasets.

For the three datasets, we treat the source tweet, retweets, and replies as nodes, and the interactions between them as edges. The creation time of edges is associated with the time of retweets or replies.

[6] https://www.politifact.com/.
[7] https://www.gossipcop.com/.

Table 1
The statistics of datasets.

Statistic	Weibo	FakeNewsNet	Twitter
# of fake news	2,131	2,079	578
# of real news	2,207	2,089	569
# of users	1,309,645	45,109	29,858
Avg. time length	1577 h	1951 h	158 h
Avg. # of tweets	378	42	30
Max. # of tweets	1999	1315	323
Min. # of tweets	10	3	2

5.2. Experimental setup

As in previous research (Bian et al., 2020; Huang, Zhou, Wu, Liu, & Bin, 2020), we split the entire data randomly into 5 equal subsamples, and then conduct 5-fold cross-validation to evaluate the model performance. We consider using a pretrained Google BERT model to get word vectors (i.e., $d = d_t = d_h = 768$) (Devlin, Chang, Lee, & Toutanova, 2019). We set d_g as 64. The head number of self-attention function in TGAT is set as 2, which follows the default settings of Rossi et al. (2020) and Xu, and Ruan et al. (2020). The hyper-parameter λ is set as $5e^{-4}$. We set the number of temporal neighbor nodes and epochs as $b = 10$ and 200, respectively. The learning rate is set to $1e^{-5}$. We choose accuracy, precision, recall, and the F_1 score as evaluation metrics, which are widely adopted in related areas (Shu, Cui, Wang, Lee, & Liu, 2019; Shu et al., 2017; Song, Ning, Zhang, & Wu, 2021).

5.3. Baseline approaches

We make comparisons with a series of baseline fake news detection methods:

- DTC (Castillo, Mendoza, & Poblete, 2011): A decision-tree-based method that employs various handcrafted features to classify fake news.
- SVM-RBF (Yang, Liu, Yu, & Yang, 2012): A support vector machine (SVM)-based method with radial basis function (RBF) kernel, which employs a series of statistics features from the tweets to identify fake news.
- SVM-TS (Ma et al., 2015): A linear SVM-based classifier that leverages time series modeling techniques to capture the temporal characteristics.
- RvNN (Ma et al., 2018b): A rumor classification method based on tree-structured recursive neural networks integrate the text content and propagation structure features using GRU units.
- StA-HiTPLAN (Khoo, Chieu, Qian, & Jiang, 2020): A transformer-networks-based fake news detection methods that incorporates time delay and propagation structure information to model long distance interactions between tweets.
- GAT (Veličković et al., 2018): It is a state-of-the-art representation learning framework but cannot utilize temporal information. The layer of GAT is set to 2, and the dimension of hidden state and output features are 768 and 64. Its output is fed into an FFN layer and a softmax layer to make the final prediction.
- GCN (Kipf & Welling, 2017): is similar to GAT. The parameter settings are same to GAT.
- VAE-GCN (Lin, Zhang, & Fu, 2020): A Variational Graph Autoencoder (VGAE)-style fake news detection methods based on GCN.
- BiGCN (Bian et al., 2020): A static graph-based fake news detection model that utilizes top-down and bottom-up GCN to learn the patterns and structures of news diffusion.
- STS-NN (Huang, & Zhou et al., 2020): A propagation-based fake news detection framework based on deep spatial temporal neural network.

DTC, SVM-TS, and SVM-RBF are *content-based methods*, and they feed the hand-engineered features to a traditional machine learning model to verify the given news. The others are *propagation-based methods* integrating both network structure and content semantics information. To make fair comparisons, the baseline methods use pre-trained Bert word vectors (Devlin et al., 2019), rather than TF-IDF values. The reason why we did not choose TF-IDF features are that the goal of this paper is to build a streaming fake news detection model, thus taking TF-IDF features may limit the ability of the model to scale with the volume of newly emerged events.

5.4. Results and analysis

In this subsection, we present comparisons against several state-of-art fake news detection methods to demonstrate the effectiveness of the proposed method. The news classification results of different methods on three datasets are shown in Tables 2–4. From these tables, we can yield several insights as follows.

- Generally, we can observe that the TGNF outperforms all baselines in terms of accuracy and F_1 score across various benchmark datasets with statistical significance, which shows the importance of temporal propagation information in verifying the authenticity of news.

Table 2
News classification results on Weibo dataset.

Method	Accuracy	Fake news			Real news		
		Precision	Recall	F_1 Score	Precision	Recall	F_1 Score
DTC[a]	0.809	0.806	0.813	0.810	0.812	0.806	0.809
SVM-RBF[a]	0.823	0.824	0.820	0.822	0.821	0.825	0.823
SVM-TS[a]	0.859	0.825	0.891	0.850	0.871	0.818	0.836
RvNN[b]	0.896	0.904	0.883	0.893	0.889	0.909	0.899
StA-HiTPLAN[b]	0.870	0.869	0.866	0.867	0.871	0.874	0.872
GAT[b]	0.931	0.924	0.937	0.931	0.939	0.926	0.932
GCN[b]	0.932	0.923	0.940	0.931	0.941	0.924	0.933
VAE-GCN[b]	0.906	0.907	0.902	0.904	0.906	0.911	0.908
BiGCN[b]	0.933	0.928	0.939	0.930	0.940	0.929	0.930
STS-NN[b]	0.912	0.912	0.908	0.910	0.911	0.915	0.913
TGNF[b]	**0.968**[c]	**0.962**[c]	**0.975**[c]	**0.969**[c]	**0.974**[c]	**0.960**[c]	**0.967**[c]

[a]Are content-based methods.
[b]Are propagation-based methods.
[c]Denotes the test of statistical significance $p < 0.01$.

Table 3
News classification results on FakeNewsNet dataset.

Method	Accuracy	Fake news			Real news		
		Precision	Recall	F_1 Score	Precision	Recall	F_1 Score
DTC[a]	0.782	0.780	0.783	0.782	0.783	0.780	0.781
SVM-RBF[a]	0.788	0.786	0.789	0.787	0.789	0.786	0.788
SVM-TS[a]	0.811	0.808	0.796	0.791	0.828	0.820	0.809
RvNN[b]	0.828	0.827	0.796	0.801	0.818	0.857	0.829
StA-HiTPLAN[b]	0.800	0.802	0.794	0.798	0.797	0.805	0.801
GAT[b]	0.885	0.886	0.883	0.884	0.884	0.887	0.885
GCN[b]	0.873	0.872	0.874	0.873	0.874	0.873	0.873
VAE-GCN[b]	0.865	0.865	0.863	0.864	0.864	0.866	0.865
BiGCN[b]	0.889	0.890	0.888	0.889	0.888	0.891	0.890
STS-NN[b]	0.858	0.867	0.847	0.857	0.848	0.868	0.858
TGNF[b]	**0.935**[c]	**0.937**[c]	**0.932**[c]	**0.935**[c]	**0.933**[c]	**0.928**[c]	**0.931**[c]

[a]Are content-based methods.
[b]Are propagation-based methods.
[c]Denotes the test of statistical significance $p < 0.01$.

- Deep learning based methods perform better than those methods based traditional machine learning. This is due to the fact that, deep learning algorithms can capture more complex patterns automatically than the methods using hand-craft features, and then can contribute to the prediction.
- Noticeably, though BiGCN shows better performance in three datasets, there is no clear winner among all the baseline approaches. However, one thing is certain is that these models based on GNN often achieve competitive results to traditional deep learning algorithms across three datasets, despite being unknown to temporal information. This is an indication that the propagation-based models have better robustness.
- We see that a slight decay in performance for all the models on Twitter dataset. One possible reason is that there are a large number of retweets in the Twitter dataset, and their node representation is the same as the source node. Strikingly, as our model can capture the temporal propagation information, it still outperforms baseline approaches by a significant margin, and achieve relative improvement of 5.8%, 5.7% in terms of Accuracy and F_1 on Twitter dataset, comparing with GAT. These results further demonstrate the necessity of modeling the temporal interaction information to discern between true and fake news.

5.5. Ablation study

In this subsection, we conduct experiments to comprehend the effect of the key modules on TGNF. More specifically, we make comparisons with the following variants of the TGNF by removing some modules in the model:

- TGNF w/o TDN: In this variant, we remove the TDN module from the model in the model training stage.
- TGNF w/o GCL+TDN: In this variant, we do not use the GCL module and TDN module in the learning phase of the model, and only consider news predictor in section 4.4.

The performance of these variants are summarized in Tables 5–7. We can make the following observations:

Table 4
News classification results on Twitter dataset.

Method	Accuracy	Fake news			Real news		
		Precision	Recall	F_1 Score	Precision	Recall	F_1 Score
DTC[a]	0.704	0.717	0.683	0.699	0.693	0.726	0.709
SVM-RBF[a]	0.732	0.740	0.724	0.731	0.725	0.741	0.733
SVM-TS[a]	0.707	0.715	0.698	0.706	0.700	0.717	0.709
RvNN[b]	0.805	0.818	0.788	0.803	0.793	0.822	0.807
StA-HiTPLAN[b]	0.780	0.777	0.783	0.780	0.782	0.776	0.779
GAT[b]	0.865	0.879	0.849	0.864	0.852	0.882	0.866
GCN[b]	0.858	0.860	0.855	0.858	0.857	0.861	0.859
VAE-GCN[b]	0.841	0.847	0.836	0.841	0.836	0.847	0.841
BiGCN[b]	0.864	0.867	0.862	0.865	0.861	0.866	0.863
STS-NN[b]	0.834	0.838	0.829	0.834	0.829	0.838	0.833
TGNF[b]	**0.923**[c]	**0.932**[c]	**0.914**[c]	**0.923**[c]	**0.914**[c]	**0.932**[c]	**0.923**[c]

[a] Are content-based methods.
[b] Are propagation-based methods.
[c] Denotes the test of statistical significance $p < 0.01$.

Table 5
Ablation study results on Weibo dataset.

Method	Accuracy	Fake news			Real news		
		Precision	Recall	F_1 Score	Precision	Recall	F_1 Score
w/o TDN	0.959	0.954	0.963	0.958	0.963	0.955	0.959
w/o GCL+TDN	0.947	0.936	0.958	0.947	0.958	0.937	0.947
TGNF	**0.968**	**0.962**	**0.975**	**0.969**	**0.974**	**0.960**	**0.967**

Table 6
Ablation study results on FakeNewsNet dataset.

Method	Accuracy	Fake news			Real news		
		Precision	Recall	F_1 Score	Precision	Recall	F_1 Score
w/o TDN	0.930	0.928	0.932	0.930	0.933	0.928	0.931
w/o GCL+TDN	0.917	0.918	0.916	0.917	0.916	0.919	0.917
TGNF	**0.935**	**0.937**	**0.932**	**0.935**	**0.933**	**0.928**	**0.931**

Table 7
Ablation study results on Twitter dataset.

Method	Accuracy	Fake news			Real news		
		Precision	Recall	F_1 Score	Precision	Recall	F_1 Score
w/o TDN	0.916	0.916	0.918	0.917	0.917	0.915	0.916
w/o GCL+TDN	0.908	0.911	0.905	0.908	0.904	0.910	0.907
TGNF	**0.923**	**0.932**	**0.914**	**0.923**	**0.914**	**0.932**	**0.923**

- As we can see from the results, TGNF outperforms its variants without the GCL module and TDN module, and removing them will reduce the performance of news classification.
- By comparing the results of the TGNF and w/o TDN, we find that paying attention to the variational information between interactions can benefit the model's performance.
- TGNF outperforms w/o GCL+TDN across three datasets, which shows that it is necessary to update the information of the influenced nodes.

In addition, We conduct experiments on the three datasets to analyze the impact of loss balancing parameter λ in Eq. (28) on accuracy. We choose various λ from the range $[5e^{-1}, 5e^{-2}, 5e^{-3}, 5e^{-4}, 5e^{-5}]$. In Fig. 5, we reported accuracy results of the TGNF without considering GCL. From Fig. 5, we can see that TDN is parameter sensitive, and can benefit prediction accuracy only if taking the appropriate parameters settings.

5.6. Early fake news detection

In this section, we evaluate the performance of our method and baseline approaches on early fake news detection. Identifying fake news at the early stage of diffusion can prevent further spreading of fake news, and help to mitigate its negative effect on society. We consider the propagation time or the released time of a certain tweet in a piece of news as the detection deadline,

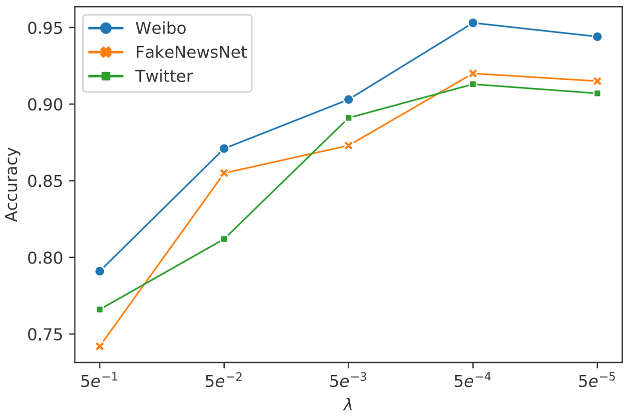


Fig. 5. Impact of λ on accuracy.

which means that the tweets published after the deadline are invisible. In this experiment, we compare different detection models by varying check time stamps in the range $\{0, 10, 20, 30, 40, 50, 60, 70, 80\}$ minutes. Fig. 6 depicts all methods' Accuracy per time point on three datasets. From Fig. 6, we can see that TGNF consistently outperforms other baselines in detecting fake news at early stage across all datasets. In all cases, their early detection Accuracy grows quickly at the early stage of propagation. However, we find that the performance of our model demonstrates obvious advantage as time goes on.

6. Conclusions

We study the problem of temporal propagation-based fake news detection task. To solve this problem, we introduce a novel fake news detection architecture named TGNF for temporal news propagation graphs in this paper. Specifically, by modeling the node's temporal interaction events, our model can capture dynamic evolution patterns of news propagation from the perspective of continuous time. We conduct extensive experiments on three real-world datasets and the experimental results demonstrate the effectiveness of the proposed framework. Our results also show that modeling and incorporating the temporal propagation information of online social media news can benefit the fake news detection task. However, there are several drawbacks to our model. First, it is difficult to propagate information to the neighbor nodes per interaction in time. Second, TGNF can only process several tweets for a piece of news each time, which means that TGNF takes a longer running time than baseline approaches (see Appendix B). In future work, we will investigate reasonable solutions to these problems. Another interesting future direction is exploring temporal heterogeneous graphs to incorporate the information social media users.

CRediT authorship contribution statement

Chenguang Song: Conceptualization, Methodology, Software, Validation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Kai Shu:** Data curation, Writing - review & editing. **Bin Wu:** Supervision.

Acknowledgments

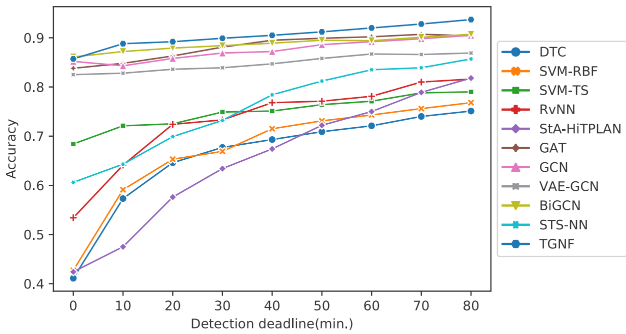
This work is supported by the NSFC-General Technology Basic Research Joint Funds (grant no. U1936220) and the National Natural Science Foundation of China (grant no. 61972047).

Appendix A. Graph convolutional networks

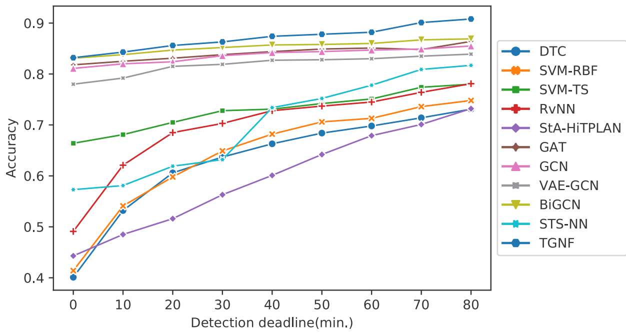
In recent years, deep learning techniques for non-Euclidean domain have gained great progress. Graph convolutional networks (GCN) are an extension of CNN on graph data, and play a representative role in combining deep learning techniques with graph data (Defferrard, Bresson, & Vandergheynst, 2016). The emergence of GCN has greatly promoted the development of the task of applying deep neural networks to graph data. For a multi-layer GCN, the layer-wise propagation rule can be formulated as follows (Kipf & Welling, 2017).

$$\mathbf{H}^{l+1} = \sigma(\tilde{\mathbf{D}}^{\frac{1}{2}} \tilde{\mathcal{A}} \tilde{\mathbf{D}}^{\frac{1}{2}} \mathbf{H}^{(l)} \mathbf{W}^{(l)}) \tag{A.1}$$

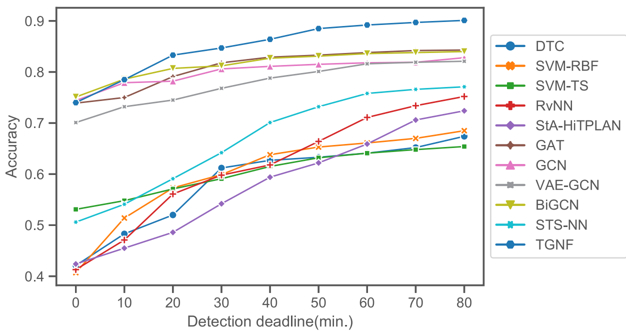
where $\mathcal{A}$ is the adjacency matrix of the graph $\mathcal{G}$, $\mathbf{I}_{\mathcal{N}}$ is the identity matrix, $\tilde{\mathcal{A}} = \mathcal{A} + \mathbf{I}_{\mathcal{N}}$ is the adjacency matrix of the graph $\mathcal{G}$ with self-connections, $\mathcal{N}$ is the number of nodes in the graph $\mathcal{G}$, $\tilde{\mathbf{D}}_{ii} = \sum_j \tilde{\mathcal{A}}_{ij}$ is the laplacian matrix of $\mathcal{G}$, $\mathbf{H}^{(l)} \in \mathbb{R}^{\mathcal{N} \times d_l}$ denotes the hidden feature matrix output by the $(l-1)$th graph convolutional layer, $\mathbf{W}^{(l)} \in \mathbb{R}^{d_l \times d_{(l+1)}}$ is a trainable parameter matrix, and $\sigma(\cdot)$ is the ReLU activation function.



(a)



(b)



(c)

Fig. 6. (a) Results of early fake news detection on Weibo dataset; (b) Results of early fake news detection on FakeNewsNet dataset; (c) Results of early fake news detection on Twitter dataset.

Table B.8
Comparison of average running time each epoch among some baselines (min).

Method	Weibo	FakeNewsNet	Twitter
GAT	1.1	0.7	0.3
GCN	0.7	0.6	0.2
BiGCN	2.9	1.5	0.8
TGNF	54.2	5.4	12.4

Appendix B. Comparison of the execution time

The static graph-based methods only need to process one completed news propagation graph each time. However, because TGNF is a temporal evolving graph-based method, temporal embedding network have to read one or a limited number of tweets once time for a piece of news, and undoubtedly spend more execution time in model running. Table B.8 compares the average running time of some models on three real world datasets in one epoch.

The reason why we choose these methods is that they are all GNN-based methods and show better performance. All experiments are conducted on GeForce RTX 2080Ti GPU. From the comparison, we observe that TGNF consistently shows a longer running time for each epoch across all the datasets. This is indeed a major drawback of the TGNF. Improving temporal update function in temporal embedding network is a feasible way to mitigate this problem.

References

Agichtein, E., Castillo, C., Donato, D., Gionis, A., & Mishne, G. (2008). Finding high-quality content in social media. In *Proceedings of the 2008 International Conference on Web Search and Data Mining* (pp. 183–194). ACM, http://dx.doi.org/10.1145/1341531.1341557.

Ajao, O., Bhowmik, D., & Zargari, S. (2018). Fake news identification on Twitter with hybrid CNN and RNN models. In *Proceedings of the 9th international conference on social media and society* (pp. 226–230). ACM, http://dx.doi.org/10.1145/3217804.3217917.

Alkhodair, S. A., Ding, S. H., Fung, B. C., & Liu, J. (2020). Detecting breaking news rumors of emerging topics in social media. *Information Processing & Management*, *57*(2), Article 102018. http://dx.doi.org/10.1016/j.ipm.2019.02.016.

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, *31*(2), 211–236. http://dx.doi.org/10.1257/jep.31.2.211.

Bian, T., Xiao, X., Xu, T., Zhao, P., Huang, W., Rong, Y., et al. (2020). Rumor detection on social media with bi-directional graph convolutional networks. In *The thirty-fourth AAAI conference on artificial intelligence* (pp. 549–556). AAAI Press.

Bondielli, A., & Marcelloni, F. (2019). A survey on fake news and rumour detection techniques. *Information Sciences*, *497*, 38–55. http://dx.doi.org/10.1016/j.ins.2019.05.035.

Cao, J., Guo, J., Li, X., Jin, Z., Guo, H., & Li, J. (2018). Automatic rumor detection on microblogs: a survey. arXiv:1807.03505.

Castillo, C., Mendoza, M., & Poblete, B. (2011). Information credibility on Twitter. In *Proceedings of the 20th international conference on world wide web* (pp. 675–684). ACM, http://dx.doi.org/10.1145/1963405.1963500.

Castillo, C., Mendoza, M., & Poblete, B. (2013). Predicting information credibility in time-sensitive social media. In P. T. M. Gayo-Avello Eni Mustafaraj, Markus Strohmaier, Harald Schoen and Peter Gloor, Daniel (Ed.), *Internet Research*, *23*(5), 560–588. http://dx.doi.org/10.1108/IntR-05-2012-0095.

Chen, T., Li, X., Yin, H., & Zhang, J. (2018). Call attention to rumors: deep attention based recurrent neural networks for early rumor detection. In *Proceedings of the 22nd pacific-asia conference on knowledge discovery and data mining* (pp. 40–52). Springer International Publishing, http://dx.doi.org/10.1007/978-3-030-04503-6_4.

Chen, Y., Sui, J., Hu, L., & Gong, W. (2019). Attention-residual network with CNN for rumor detection. In *Proceedings of the 28th ACM international conference on information and knowledge management* (pp. 1121–1130). ACM, http://dx.doi.org/10.1145/3357384.3357950.

Cheng, M., Nazarian, S., & Bogdan, P. (2020). Vroc: Variational autoencoder-aided multi-task rumor classifier based on text. In *Proceedings of the web conference 2020* (pp. 2892–2898). ACM, http://dx.doi.org/10.1145/3366423.3380054.

Chow, S.-N., Ye, X., Zha, H., & Zhou, H. (2018). Influence prediction for continuous-time information propagation on networks. *Networks & Heterogeneous Media*, *13*(4), 567–583.

Cui, L., Seo, H., Tabar, M., Ma, F., Wang, S., & Lee, D. (2020). DETERRENT: Knowledge guided graph attention network for detecting healthcare misinformation. In *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining* (pp. 492–502). ACM, http://dx.doi.org/10.1145/3394486.3403092.

Cui, L., Wang, S., & Lee, D. (2019). SAME: Sentiment-aware multi-modal embedding for detecting fake news. In *Proceedings of the 2019 IEEE/ACM international conference on advances in social networks analysis and mining* (pp. 41–48). ACM, http://dx.doi.org/10.1145/3341161.3342894.

Defferrard, M., Bresson, X., & Vandergheynst, P. (2016). Convolutional neural networks on graphs with fast localized spectral filtering. In *Proceedings of the 30th international conference on neural information processing systems* (pp. 3844–3852). Curran Associates Inc..

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., et al. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, *113*(3), 554–559. http://dx.doi.org/10.1073/pnas.1517441113.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the north american chapter of the association for computational linguistics* (pp. 4171–4186). Association for Computational Linguistics.

Dun, Y., Tu, K., Chen, C., Hou, C., & Yuan, X. (2021). KAN: Knowledge-aware attention network for fake news detection. In *Proceedings of the 35th international joint conference on artificial intelligence*. AAAI Press.

Dungs, S., Aker, A., Fuhr, N., & Bontcheva, K. (2018). Can rumour stance alone predict veracity? In *Proceedings of the 27th international conference on computational linguistics* (pp. 3360–3370). ACL.

Gergely, P., Albert-László, B., & Tamás, V. (2007). Quantifying social group evolution. *Nature*, *446*, 3–6. http://dx.doi.org/10.1038/nature05670.

Goldani, M. H., Momtazi, S., & Safabakhsh, R. (2021). Detecting fake news with capsule neural networks. *Applied Soft Computing*, *101*, Article 106991. http://dx.doi.org/10.1016/j.asoc.2020.106991.

Goldani, M. H., Safabakhsh, R., & Momtazi, S. (2021). Convolutional neural network with margin loss for fake news detection. *Information Processing & Management*, *58*(1), Article 102418. http://dx.doi.org/10.1016/j.ipm.2020.102418.

Gomez-Rodriguez, M., & Schölkopf, B. (2012). Influence maximization in continuous time diffusion networks. In *Proceedings of the 29th international coference on international conference on machine learning* (pp. 579–586). Madison, WI, USA: Omnipress.

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al. (2014). Generative adversarial nets. In *NIPS*.

Guo, H., Cao, J., Zhang, Y., Guo, J., & Li, J. (2018). Rumor detection with hierarchical social attention network. In *Proceedings of the 27th ACM international conference on information and knowledge management* (pp. 943–951). ACM, http://dx.doi.org/10.1145/3269206.3271709.

Guo, B., Ding, Y., Yao, L., Liang, Y., & Yu, Z. (2020). The future of false information detection on social media: New perspectives and trends. *ACM Computing Surveys*, *53*(4), http://dx.doi.org/10.1145/3393880.

Hamilton, W. L., Ying, R., & Leskovec, J. (2017). Inductive representation learning on large graphs. In *NIPS*.

Han, Y., Huang, G., Song, S., Yang, L., Wang, H., & Wang, Y. (2021). Dynamic neural networks: A survey. arXiv:2102.04906.

Holme, P. (2015). Modern temporal network theory: a colloquium. *The European Physical Journal B*, *88*(9), 1–30.

Holme, P., & Saramäki, J. (2012). Temporal networks. *Physics Reports*, *519*(3), 97–125. http://dx.doi.org/10.1016/j.physrep.2012.03.001.

Huang, Q., Yu, J., Wu, J., & Wang, B. (2020). Heterogeneous graph attention networks for early detection of rumors on Twitter. In *2020 international joint conference on neural networks* (pp. 1–8). IEEE, http://dx.doi.org/10.1109/IJCNN48605.2020.9207582.

Huang, Q., Zhou, C., Wu, J., Liu, L., & Wang, B. (2020). Deep spatial–temporal structure learning for rumor detection on Twitter. *Neural Computing and Applications*, 1–11.

Huang, Q., Zhou, C., Wu, J., Wang, M., & Wang, B. (2019). Deep structure learning for rumor detection on Twitter. In *2019 International joint conference on neural networks* (pp. 1–8). http://dx.doi.org/10.1109/IJCNN.2019.8852468.

Jian, L., Li, J., & Liu, H. (2018). Toward online node classification on streaming networks. *Data Mining and Knowledge Discovery*, *32*(1), 231–257. http://dx.doi.org/10.1007/s10618-017-0533-y.

Jin, Z., Cao, J., Guo, H., Zhang, Y., & Luo, J. (2017). Multimodal fusion with recurrent neural networks for rumor detection on microblogs. In *Proceedings of the 25th ACM international conference on multimedia* (pp. 795–816). ACM, http://dx.doi.org/10.1145/3123266.3123454.

Kazemi, S. M., Goel, R., Eghbali, S., Ramanan, J., Sahota, J., Thakur, S., et al. (2019). Time2Vec: Learning a vector representation of time. arXiv:1907.05321.

Kazemi, S. M., Goel, R., Jain, K., Kobyzev, I., Sethi, A., Forsyth, P., et al. (2020). Representation learning for dynamic graphs: A survey. *Journal of Machine Learning Research*, *21*(70), 1–73.

Khattar, D., Goud, J. S., Gupta, M., & Varma, V. (2019). MVAE: multimodal variational autoencoder for fake news detection. In *Proceedings of the 28th international conference on world wide web* (pp. 2915–2921). ACM, http://dx.doi.org/10.1145/3308558.3313552.

Khoo, L. M. S., Chieu, H. L., Qian, Z., & Jiang, J. (2020). Interpretable rumor detection in microblogs by attending to user interactions. In *The thirty-fourth AAAI conference on artificial intelligence* (pp. 8783–8790). AAAI Press.

Kipf, T. N., & Welling, M. (2017). Semi-supervised classification with graph convolutional networks. In *International conference on learning representations*.

Kleinberg, J. (2006). Temporal dynamics of on-line information streams. In *Data stream management: processing high-speed data*. Citeseer.

Kochkina, E., Liakata, M., & Zubiaga, A. (2018). All-in-one: multi-task learning for rumour verification. In *Proceedings of the 27th international conference on computational linguistics* (pp. 3402–3413). ACL.

Kumar, S., & Shah, N. (2018). False information on web and social media: A survey. arXiv:1804.08559.

Kumar, S., Zhang, X., & Leskovec, J. (2019). Predicting dynamic embedding trajectory in temporal interaction networks. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining* (pp. 1269–1278). ACM, http://dx.doi.org/10.1145/3292500.3330895.

Li, Y., Tam, D. S. H., Xie, S., Liu, X., Ying, Q. F., Lau, W. C., et al. (2020). GTEA: Representation learning for temporal interaction graphs via edge aggregation. arXiv:2009.05266.

Li, Z., Zhang, Q., Du, X., Ma, Y., & Wang, S. (2021). Social media rumor refutation effectiveness: Evaluation, modelling and enhancement. *Information Processing & Management*, *58*(1), Article 102420. http://dx.doi.org/10.1016/j.ipm.2020.102420.

Lin, H., Zhang, X., & Fu, X. (2020). A graph convolutional encoder and decoder model for rumor detection. In *2020 IEEE 7th international conference on data science and advanced analytics* (pp. 300–306). http://dx.doi.org/10.1109/DSAA49011.2020.00043.

Liu, Y., Jin, X., & Shen, H. (2019). Towards early identification of online rumors based on long short-term memory networks. *Information Processing & Management*, *56*(4), 1457–1467. http://dx.doi.org/10.1016/j.ipm.2018.11.003.

Liu, Y., & Wu, Y. B. (2018). Early detection of fake news on social media through propagation path classification with recurrent and convolutional networks. In S. A. McIlraith, & K. Q. Weinberger (Eds.), *Proceedings of the thirty-second AAAI conference on artificial intelligence* (pp. 354–361). AAAI Press.

Liu, Y., & Wu, Y.-F. B. (2020). FNED: A deep network for fake news early detection on social media. *ACM Transactions on Information Systems*, *38*(3), http://dx.doi.org/10.1145/3386253.

Lu, Y.-J., & Li, C.-T. (2020). GCAN: Graph-aware co-attention networks for explainable fake news detection on social media. In *Proceedings of the 58th annual meeting of the association for computational linguistics* (pp. 505–514). ACL, http://dx.doi.org/10.18653/v1/2020.acl-main.48, Online.

Lu, Y., Wang, X., Shi, C., Yu, P. S., & Ye, Y. (2019). Temporal network embedding with micro- and macro-dynamics. In *Proceedings of the 28th ACM international conference on information and knowledge management* (pp. 469–478). ACM, http://dx.doi.org/10.1145/3357384.3357943.

Lukasik, M., Cohn, T., & Bontcheva, K. (2015). Point process modelling of rumour dynamics in social media. In *Proceedings of the 53rd annual meeting of the association for computational linguistics and the 7th international joint conference on natural language processing* (pp. 518–523). ACL, http://dx.doi.org/10.3115/v1/P15-2085.

Ma, J., Gao, W., Mitra, P., Kwon, S., Jansen, B. J., Wong, K.-F., et al. (2016). Detecting rumors from microblogs with recurrent neural networks. In *Proceedings of the twenty-fifth international joint conference on artificial intelligence* (pp. 3818–3824). AAAI Press.

Ma, J., Gao, W., Wei, Z., Lu, Y., & Wong, K.-F. (2015). Detect rumors using time series of social context information on microblogging websites. In *Proceedings of the 24th ACM International on conference on information and knowledge management* (pp. 1751–1754). ACL, http://dx.doi.org/10.1145/2806416.2806607.

Ma, J., Gao, W., & Wong, K.-F. (2017). Detect rumors in microblog posts using propagation structure via kernel learning. In *Proceedings of the 55th annual meeting of the association for computational linguistics* (pp. 708–717). ACL, http://dx.doi.org/10.18653/v1/P17-1066.

Ma, J., Gao, W., & Wong, K.-F. (2018a). Detect rumor and stance jointly by neural multi-task learning. In *Companion proceedings of the the web conference 2018* (pp. 585–593). International World Wide Web Conferences Steering Committee, http://dx.doi.org/10.1145/3184558.3188729.

Ma, J., Gao, W., & Wong, K.-F. (2018b). Rumor detection on twitter with tree-structured recursive neural networks. In *Proceedings of the 56th annual meeting of the association for computational linguistics* (pp. 1980–1989). Melbourne, Australia: ACL, http://dx.doi.org/10.18653/v1/P18-1184.

Ma, J., Gao, W., & Wong, K.-F. (2019). Detect rumors on twitter by promoting information campaigns with generative adversarial learning. In *Proceedings of the 28th international conference on world wide web* (pp. 3049–3055). ACM, http://dx.doi.org/10.1145/3308558.3313741.

Ma, Y., Guo, Z., Ren, Z., Tang, J., & Yin, D. (2020). Streaming graph neural networks. In *Proceedings of the 43rd International ACM SIGIR conference on research and development in information retrieval* (pp. 719–728). ACM, http://dx.doi.org/10.1145/3397271.3401092.

Manessi, F., Rozza, A., & Manzo, M. (2020). Dynamic graph convolutional networks. *Pattern Recognition*, *97*, Article 107000. http://dx.doi.org/10.1016/j.patcog.2019.107000.

Meel, P., & Vishwakarma, D. K. (2020). Fake news, rumor, information pollution in social media and web: A contemporary survey of state-of-the-arts, challenges and opportunities. *Expert Systems with Applications*, *153*, Article 112986. http://dx.doi.org/10.1016/j.eswa.2019.112986.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In *Proceedings of the 26th international conference on neural information processing systems* (pp. 3111–3119). Curran Associates Inc., http://dx.doi.org/10.5555/2999792.2999959.

Monti, F., Frasca, F., Eynard, D., Mannion, D., & Bronstein, M. M. (2019). Fake news detection on social media using geometric deep learning. arXiv:1902.06673.

Nguyen, V. H., Sugiyama, K., Nakov, P., & Kan, M. Y. (2020). FANG: Leveraging social context for fake news detection using graph representation. In *Proceedings of the 29th ACM international conference on information & knowledge management* (pp. 1165–1174). ACM, http://dx.doi.org/10.1145/3340531.3412046.

Qi, P., Cao, J., Yang, T., Guo, J., & Li, J. (2019). Exploiting multi-domain visual information for fake news detection. In *2019 IEEE international conference on data mining* (pp. 518–527).

Rossi, E., Chamberlain, B., Frasca, F., Eynard, D., Monti, F., & Bronstein, M. (2020). Temporal graph networks for deep learning on dynamic graphs. In *ICML 2020 workshop on graph representation learning*.

Rubin, V. L., Chen, Y., & Conroy, N. J. (2015). Deception detection for news: three types of fakes. In *Proceedings of the 78th ASIS&T annual meeting: information science with impact: research in and for the community* (pp. 83:1–83:4). American Society for Information Science.

Ruchansky, N., Seo, S., & Liu, Y. (2017). CSI: A hybrid deep model for fake news detection. In *Proceedings of the 2017 ACM on conference on information and knowledge management* (pp. 797–806). ACM, http://dx.doi.org/10.1145/3132847.3132877.

Saito, K., Kimura, M., Ohara, K., & Motoda, H. (2009). Learning continuous-time information diffusion model for social behavioral data analysis. In *Proceedings of the 1st asian conference on machine learning: advances in machine learning* (pp. 322–337). Berlin, Heidelberg: Springer-Verlag, http://dx.doi.org/10.1007/978-3-642-05224-8_25.

Sankar, A., Wu, Y., Gou, L., Zhang, W., & Yang, H. (2020). Dysat: Deep neural representation learning on dynamic graphs via self-attention networks. In *Proceedings of the 13th international conference on web search and data mining* (pp. 519–527). New York, NY, USA: ACM, http://dx.doi.org/10.1145/3336191.3371845.

Sejeong, K., Meeyoung, C., Kyomin, J., Wei, C., & Yajun, W. (2013). Prominent features of rumor propagation in online social media. In *IEEE 13th international conference on data mining* (pp. 1103–1108). IEEE, http://dx.doi.org/10.1109/ICDM.2013.61.

Sharma, K., Qian, F., Jiang, H., Ruchansky, N., Zhang, M., & Liu, Y. (2019). Combating fake news: a survey on identification and mitigation techniques. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *10*(3), 21:1–21:42. http://dx.doi.org/10.1145/3305260.

Shu, K., Cui, L., Wang, S., Lee, D., & Liu, H. (2019). dEFEND: explainable fake news detection. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining* (pp. 395–405). ACM, http://dx.doi.org/10.1145/3292500.3330935.

Shu, K., Mahudeswaran, D., Wang, S., Lee, D., & Liu, H. (2020). FakeNewsNet: A data repository with news content, social context, and spatiotemporal information for studying fake news on social media. *Big Data*, *8*(3), 171–188. http://dx.doi.org/10.1089/big.2020.0062.

Shu, K., Mahudeswaran, D., Wang, S., & Liu, H. (2020). Hierarchical propagation networks for fake news detection: Investigation and exploitation. In *Proceedings of the international AAAI conference on web and social media* (pp. 626–637). Vol. 14 (1).

Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: a data mining perspective. *ACM SIGKDD Explorations Newsletter*, *19*(1), 22–36. http://dx.doi.org/10.1145/3137597.3137600.

Shu, K., Zheng, G., Li, Y., Mukherjee, S., Awadallah, A. H., Ruston, S., et al. (2020). Early detection of fake news with multi-source weak social supervision. In *Proceedings of the european conference on machine learning and principles and practice of knowledge discovery in databases 2020*.

Si, M., Cui, L., Guo, W., Li, Q., Liu, L., Lu, X., et al. (2020). A comparative analysis for spatio-temporal spreading patterns of emergency news. *Scientific Reports*, *10*(1), 1–13.

Silva, A., Han, Y., Luo, L., Karunasekera, S., & Leckie, C. (2020). Embedding partial propagation network for fake news early detection. In *Proceedings of the CIKM 2020 workshops (5th international workshop on mining actionable insights from social networks)*. Vol. 2699.

Silva, A., Luo, L., Karunasekera, S., & Leckie, C. (2021). Embracing domain differences in fake news: Cross-domain fake news detection using multimodal data. In *Proceedings of the 35th international joint conference on artificial intelligence*. AAAI Press.

Song, C., Ning, N., Zhang, Y., & Wu, B. (2021). A multimodal fake news detection model based on crossmodal attention residual and multichannel convolutional neural networks. *Information Processing & Management*, *58*(1), Article 102437. http://dx.doi.org/10.1016/j.ipm.2020.102437.

Song, C., Yang, C., Chen, H., Tu, C., Liu, Z., & Sun, M. (2019). CED: Credible early detection of social media rumors. *IEEE Transactions on Knowledge and Data Engineering*, 1. http://dx.doi.org/10.1109/TKDE.2019.2961675.

Tang, J. K. (2012). *Temporal network metrics and their application to real world networks* (Ph.D. thesis), UK: University of Cambridge.

Tian, T., Liu, Y., Yang, X., Lyu, Y., Zhang, X., & Fang, B. (2020). QSAN: A quantum-probability based signed attention network for explainable false information detection. In *Proceedings of the 29th ACM international conference on information & knowledge management* (pp. 1445–1454). ACM, http://dx.doi.org/10.1145/3340531.3411890.

Trivedi, R., Farajtabar, M., Biswal, P., & Zha, H. (2019). DyRep: Learning representations over dynamic graphs. In *International conference on learning representations*.

Tu, K., Chen, C., Hou, C., Yuan, J., Li, J., & Yuan, X. (2021). Rumor2vec: A rumor detection framework with joint text and propagation structure representation learning. *Information Sciences*, *560*, 137–151. http://dx.doi.org/10.1016/j.ins.2020.12.080.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). Attention is all you need. In *Advances in neural information processing systems* (pp. 5998–6008). Curran Associates, Inc..

Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., & Bengio, Y. (2018). Graph attention networks. In *International conference on learning representations*.

Vishwakarma, D. K., Varshney, D., & Yadav, A. (2019). Detection and veracity analysis of fake news via scrapping and authenticating the web search. *Cognitive Systems Research*, *58*, 217–229. http://dx.doi.org/10.1016/j.cogsys.2019.07.004.

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146–1151. http://dx.doi.org/10.1126/science.aap9559.

Vu, D. T., & Jung, J. J. (2021). Rumor detection by propagation embedding based on graph convolutional network. *International Journal of Computational Intelligence Systems*, *14*, 1053–1065. http://dx.doi.org/10.2991/ijcis.d.210304.002.

Wang, W. Y. (2017). "liar, liar pants on fire": A new benchmark dataset for fake news detection. In *Proceedings of the 55th annual meeting of the association for computational linguistics (volume 2: short papers)* (pp. 422–426). ACL, http://dx.doi.org/10.18653/v1/P17-2067.

Wang, Y., Ma, F., Jin, Z., Yuan, Y., Xun, G., Jha, K., et al. (2018). EANN: event adversarial neural networks for multi-modal fake news detection. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining* (pp. 849–857). ACM, http://dx.doi.org/10.1145/3219819.3219903.

Wei, P., Xu, N., & Mao, W. (2019). Modeling conversation structure and temporal dynamics for jointly predicting rumor stance and veracity. In *Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing* (pp. 4787–4798). ACL, http://dx.doi.org/10.18653/v1/D19-1485.

Wu, Z., Pi, D., Chen, J., Xie, M., & Cao, J. (2020). Rumor detection based on propagation graph neural network with attention mechanism. *Expert Systems with Applications*, *158*, Article 113595. http://dx.doi.org/10.1016/j.eswa.2020.113595.

Xu, D., Ruan, C., Korpeoglu, E., Kumar, S., & Achan, K. (2020). Inductive representation learning on temporal graphs. In *International conference on learning representations*.

Xu, N., Zeng, Z., & Mao, W. (2020). Reasoning with multimodal sarcastic tweets via modeling cross-modality contrast and semantic association. In *Proceedings of the 58th annual meeting of the association for computational linguistics* (pp. 3777–3786). ACL, http://dx.doi.org/10.18653/v1/2020.acl-main.349.

Yang, F., Liu, Y., Yu, X., & Yang, M. (2012). Automatic detection of rumor on sina weibo. In *Proceedings of the ACM SIGKDD workshop on mining data semantics* (pp. 13:1–13:7). ACM. http://dx.doi.org/10.1145/2350190.2350203.

Yang, X., Lyu, Y., Tian, T., Liu, Y., Liu, Y., & Zhang, X. (2020). Rumor detection on social media with graph structured adversarial learning. In *Proceedings of the twenty-ninth international joint conference on artificial intelligence* (pp. 1417–1423). International Joint Conferences on Artificial Intelligence Organization.

Yu, F., Liu, Q., Wu, S., Wang, L., & Tan, T. (2017). A convolutional approach for misinformation identification. In *Proceedings of the 26th international joint conference on artificial intelligence* (pp. 3901–3907). AAAI Press, http://dx.doi.org/10.24963/ijcai.2017/545.

Yuan, C., Ma, Q., Zhou, W., Han, J., & Hu, S. (2019). Jointly embedding the local and global relations of heterogeneous graph for rumor detection. In *2019 IEEE international conference on data mining* (pp. 796–805). http://dx.doi.org/10.1109/ICDM.2019.00090.

Yuan, C., Ma, Q., Zhou, W., Han, J., & Hu, S. (2020). *Early detection of fake news by utilizing the credibility of news, publishers, and users based on weakly supervised learning* (pp. 5444–5454). ACL, http://dx.doi.org/10.18653/v1/2020.coling-main.475.

Zannettou, S., Sirivianos, M., Blackburn, J., & Kourtellis, N. (2019). The web of false information: Rumors, fake news, hoaxes, clickbait, and various other shenanigans. *Journal of Data and Information Quality*, *11*(3), http://dx.doi.org/10.1145/3309699.

Zhang, Z., Bu, J., Ester, M., Zhang, J., Yao, C., Li, Z., et al. (2020). Learning temporal interaction graph embedding via coupled memory networks. In *Proceedings of the web conference 2020* (pp. 3049–3055). ACM, http://dx.doi.org/10.1145/3366423.3380076.

Zhang, Q., Cook, J., & Yilmaz, E. (2021). Detecting and forecasting misinformation via temporal and geometric propagation patterns. In D. Hiemstra, M. Moens, J. Mothe, R. Perego, M. Potthast, & F. Sebastiani (Eds.), *Lecture notes in computer science*: *vol. 12657*, *Advances in information retrieval - 43rd european conference on IR research* (pp. 455–462). Springer, http://dx.doi.org/10.1007/978-3-030-72240-1_48.

Zhang, H., Fang, Q., Qian, S., & Xu, C. (2019). Multi-modal knowledge-aware event memory network for social media rumor detection. In *Proceedings of the 27th ACM international conference on multimedia* (pp. 1942–1951). ACM, http://dx.doi.org/10.1145/3343031.3350850.

Zhang, X., & Ghorbani, A. A. (2020). An overview of online fake news: Characterization, detection, and discussion. *Information Processing & Management*, *57*(2), Article 102025. http://dx.doi.org/10.1016/j.ipm.2019.03.004.

Zhao, Y., Da, J., & Yan, J. (2021). Detecting health misinformation in online health communities: Incorporating behavioral features into machine learning based approaches. *Information Processing & Management*, *58*(1), Article 102390. http://dx.doi.org/10.1016/j.ipm.2020.102390.

Zhou, K., Shu, C., Li, B., & Lau, J. H. (2019). Early rumour detection. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies* (pp. 1614–1623). Minneapolis, Minnesota: Association for Computational Linguistics, http://dx.doi.org/10.18653/v1/N19-1163.

Zhou, X., Wu, J., & Zafarani, R. (2020). SAFE: Similarity-aware multi-modal fake news detection. In H. W. Lauw, R. C.-W. Wong, A. Ntoulas, E.-P. Lim, S.-K. Ng, & S. J. Pan (Eds.), *Advances in knowledge discovery and data mining* (pp. 354–367). Cham: Springer International Publishing.

Zhou, C., Xiu, H., Wang, Y., & Yu, X. (2021). Characterizing the dissemination of misinformation on social media in health emergencies: An empirical study based on COVID-19. *Information Processing & Management*, *58*(4), Article 102554. http://dx.doi.org/10.1016/j.ipm.2021.102554.

Zhou, X., & Zafarani, R. (2019). Network-based fake news detection: A pattern-driven approach. *SIGKDD Exploring Newsletter*, *21*(2), 48–60. http://dx.doi.org/10.1145/3373464.3373473.

Zhou, X., Zafarani, R., Shu, K., & Liu, H. (2019). Fake news: fundamental theories, detection strategies and challenges. In *Proceedings of the twelfth ACM international conference on web search and data mining* (pp. 836–837). ACM, http://dx.doi.org/10.1145/3289600.3291382.

Zubiaga, A., Kochkina, E., Liakata, M., Procter, R., Lukasik, M., Bontcheva, K., et al. (2018). Discourse-aware rumour stance classification in social media using sequential classifiers. *Information Processing & Management*, *54*(2), 273–290. http://dx.doi.org/10.1016/j.ipm.2017.11.009.