Detection of Fake News Through Heterogeneous Graph Interactions

Raed Alharbi, Tre' R. Jeter, and My T. Thai(✉)

University of Florida, Gainesville, FL 32603, USA
{r.alharbi,t.jeter,mythai}@ufl.edu

Abstract. Fake news is one of the most prominent forms of disinformation. Unfortunately, today's advanced social media platforms allow for the rapid transmission of fake news, which may negatively impact several aspects of human society. Despite the significant progress in detecting fake news, the focus of most current work lies in the detection based on content-based or user context-based methods. We believe that such methods suffer from two limitations: the lack of characterizing the news variation (fake news can appear in different forms via tweets, such as writing different tweets about the same news article) and news repetition (fake news is shared repeatedly via retweets); and the absence of the temporal engagement among different social interactions. Thus, we propose a novel detection framework, namely the Temporal graph Fake News Detection Framework (T-FND), that is effectively able to capture heterogeneous and repetitive characteristics of fake news behavior, resulting in better prediction performance. We empirically evaluate the effectiveness of our model on two real-world datasets, showing that our solution outperforms the state-of-the-art baseline methods

Keywords: Fake news detection · Disinformation · GNN

1 Introduction

The popularity and widespread usage of social media platforms have significantly changed the way individuals communicate and engage online, producing an ideal environment for disseminating fake news [6]. The massive transmission of fake news has started a data war in recent years, raising disinformation and fake news concerns, reducing the legitimacy of news outlets in ecosystems, and perhaps affecting readers' views on serious matters for our community [11].

Despite the remarkable improvement achieved in exploring propagation patterns of information and how they might be used in fake news detection models [10,11], we believe it is still essential to understand news propagation structures and how the news interact with each other in a social media environment to develop an effective fake news detection model. For example, on Twitter, fake news is often spread by a number of tweets about a news article [11], followed by a series of retweets, along with their temporal engagement and users' information, as represented in Fig. 1.

D. Mohaisen and T. Wies (Eds.): NETYS 2023, LNCS 14067, pp. 3–16, 2023.
https://doi.org/10.1007/978-3-031-37765-5_1

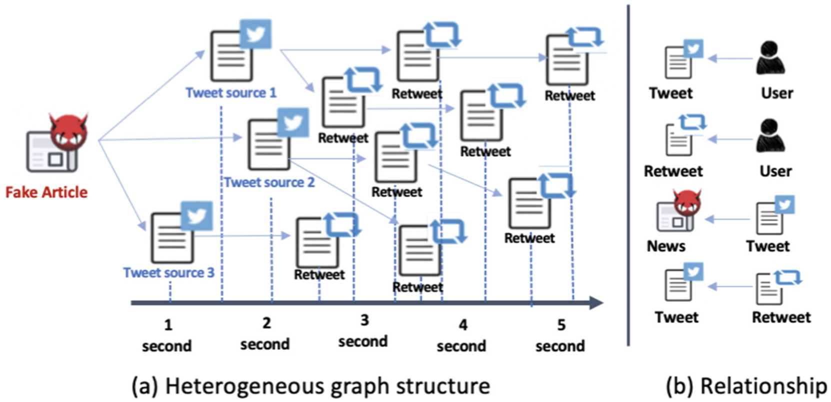

Fig. 1. Heterogeneous graph structure on Twitter. (a) A propagation example of one fake news article. (b) Different connections in the heterogeneous network between nodes.

Modeling such propagation structure is crucial to fake news detection for varied purposes. Variation and repetition are two key characteristics that can distinguish patterns of fake news patterns from those of real news. (i) *Variation*: Fake news is carried out continuously in various forms (different tweets provide different descriptions of a news article), and (ii) *Repetition*: Fake news is repeatedly shared in the form of retweets by one or more users. In the real world, news propagation networks often involve multi-levels (variation and repetition). Exploiting such news entities' interactions by analyzing the content of tweets/retweets implies that the propagation network of news on social networks might be utilized effectively to predict fake news [10]. Another key feature that might be exploited in detecting fake news is a relatively short period of time between sequential article-tweets/tweet-retweets to limit the fast spread of fake news.

However, modeling the heterogeneous propagation structure on Twitter presents a number of challenges. (i) **Sparsity.** The dynamic structure of entities' interactions, where most retweets are directed toward tweet users, results in a "hub" where the owner's degree is significantly higher than the retweets' degrees [7]. As a result, the interaction graph might be sparse. (ii) **Repetition.** Characterizing the repetition of fake news activity is challenging because we must satisfy at least two requirements concurrently: the fake news context behaviors (in different forms) and temporal engagement characteristics among different interactions.

To address these concerns, this paper proposes a **T**emporal-graph **F**ake **N**ews **D**etection framework (T-FND) that consists of three components: (1) Social context representation; (2) Temporal graph learning; and (3) A heterogeneous social network. The first part of the model aims to extract tweets and retweets representations of texts along with user information characteristics and then approximate those values features. The temporal graph in the second part is to construct a relationship interactions graph (e.g., tweet-article and retweet-tweet) by encoding those interactions' content and temporal engagement behavior. The last part selects the interaction graph with informative meta path (relationship interactions), aggregates its information representations, and then performs the fake news detection. Extensive experiments on two benchmarks demonstrate our approach's efficacy.

Organization. The remainder of the paper is structured as follows. Section 2 presents the related works. Our proposed T-FND model is introduced in Sect. 3 while the experimental evaluations are discussed in Sect. 4. Finally, Sects. 5 and 6 present the paper's limitations and conclusion, respectively.

2 Related Work

Graph-Based Fake News Detection. Graph neural networks (GNNs) have gained popularity as a reliable technique for modeling real world applications, like fake news detection [11]. The graph-based fake news detection exploits not only content and context of texts but also the interactions among various nodes in the graph, such as user-user interactions or content-content interactions [4]. Understanding the interaction patterns of fake news dissemination is critical since it provides valuable information for identifying fake news.

Many efforts have been presented to learn graph representations efficiently. The Graph Attention Network (GAT) learns feature representation of nodes by using an attention approach to distribute multiple weights for the properties of surrounding nodes [12]. The Hierarchical Attention Network (HAN) in [13] learns multiple sorts of relationships and interactions between nodes via a hierarchical attention mechanism on a heterogeneous graph, whereas the Heterogeneous Graph Neural Network (HetGNN) model aggregates and learns heterogeneous structural information on a graph to handle a variety of graph tasks [15].

Recent efforts have developed graph embeddings to transform detailed network information into organized multi-dimensional characteristics to leverage the most informative features from graph interactions. Dong *et al.* in [3] proposes Metapath2vec, a heterogeneous network representation that uses the meta path-based random walk to perform node embeddings utilizing skip-gram models. In contrast, Yun *et all* propose Graph Transformer Network (GTN), a method to learn node representation on a heterogeneous graph in which no predefined meta-paths are required [14].

Our work is related to exploring GTN for fake news detection. Characterizing interactions between tweets, retweets, or users is central to the modeling process. Yet, no attempt has been made previously to explore graph transformer network for fake news detection. Therefore, to maximize the utility of GTN, we incorporate two critical aspects: temporal features into the edge weight calculation and a fake news classifier.

Context-Based Methods. Context-based approaches are primarily concerned with features generated from user-based, post-based, and network-based data [9]. Previous efforts have relied on a wide range of contextual features with the aim of detecting fake content, such as using posts and comments [6]. Such features can help in describing the details of user behavior and the process of news propagation across time [11], providing crucial supplementary information for determining the authenticity of news articles.

Unlike conventional methods that focus exclusively on news content or one of the context-based features, we believe that effective collaboration of heterogeneous social media information, such as the content and context of tweets, retweets, user information, social media engagement with users, and temporal features, are informative to improve the detection of fake news.

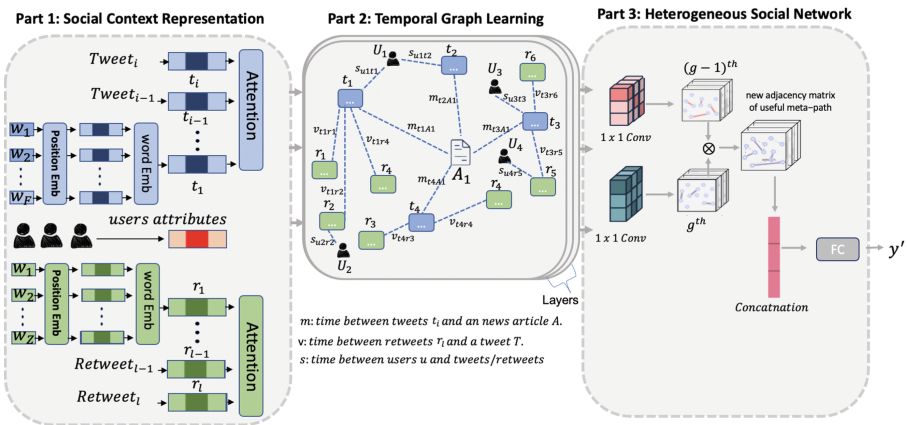

Fig. 2. Overview of T-FND. **Left.** The social context representation component hierarchically establishes the tweet sequences t_i (blue), retweet sequences r_l (green), and user attributes. **Middle.** The temporal graph learning component establishes a heterogeneous interaction graph with their temporal engagement interactions, where nodes represent tweets, retweets, users, and edges reflect the time duration between various relationships. **Right.** The heterogeneous social network component aggregates the information representation learned from graph interactions and leverages the idea of GTN to learn the new graph with a proper meta path and then perform the final classification (fake or real news). (Color figure online)

3 The Proposed Framework - T-FND

We develop a **T**emporal-graph **F**ake **N**ews **D**etection framework (T-FND) to detect fake news by incorporating heterogeneous data from various categories, including tweets, retweets, and user characteristics along with their temporal engagement features. The architecture of the framework is depicted in Fig. 2.

T-FND consists of three parts: (1) A social context component to encode (i) tweets and retweets based on their textual contents and (ii) users by considering the information reflected in their profiles. (2) A temporal graph learning component that constructs an interactions graph of different entities and models their temporal properties. The proposed graph jointly captures the diverse features, including tweets and retweets content information, user characteristics, and temporal engagement features between different interactions; and (3) A heterogeneous social network component that aggregates the information representation learned from graph interactions and leverages the idea of GTN to learn the news graph with a useful meta path to then perform the final classification (fake or real news).

3.1 Social Context Representation Component

Representing the properties of tweets, retweets and a user's historical information is essential for determining (i) user intent, which appears in the context of their utterances, and (ii) user traits such as personalities and linguistic patterns. Thus, the goal of the social context representation component in (part 1 in Fig. 2) is to extract the tweets and retweets features of texts and approximate those informative features. We begin with a description of tweets and retweets' characteristics.

Context Encoding. The variation of a news article in the form of tweets and the repeated sharing of retweets can impact large users to distribute fake news, such as making people post fake news more often. Thus, the content of tweets is the first sign of fake news detection, whereas understanding retweets characteristics is also crucial to the detection process as the sharing mechanism is commonly used in the transmission of fake news, which often includes one or more users.

To capture those representations, we divide the social context representation into three levels. The first level is segment embedding, which is used to separate sentences within a sentence pair and is subsequently used to assess whether a certain embedding belongs to a specific sentence pair. The second level employs position embedding in a sentence to define a word's position inside a sentence, such that each location has its own distinct representation (word context for a sentence). The last level is the word embedding, which converts the tweet $t_i = \{W_1, ..., W_M\}$, where M is the word size of the tweet, and the retweet $r_l = \{W_1, ..., W_Z\}$, where Z is the retweet size word, to a series of vectors with high dimensions (hidden state) $t_i = \{h_1, ..., h_M\}$ and $r_l = \{h_1, ..., h_Z\}$, respectively.

Nevertheless, the words in a tweet or retweet are not equally important in the context of fake news. Therefore, we use word-level attention to identify informative words from a tweet or retweet automatically. Hence, the tweet vector representation t of M words, and the retweet vector representation r of Z words are computed as follow:

$$t_i = \sum_{\tau=1}^{M} \alpha_{i\tau} h_{i\tau} \qquad (1) \qquad\qquad r_l = \sum_{\omega=1}^{Z} e_{l\omega} h_{l\omega} \qquad (2)$$

where $\alpha_{i\tau}$ represents the attention weights of the τ^{th} word in the tweet i and $e_{l\omega}$ represents the attention weights of the ω^{th} word in the retweet l. $h_{i\tau}$ and $h_{l\omega}$ refer to tweet and retweet words vector representations, which is obtained utilizing the (**UNI**fied pre-trained **L**anguage **M**odel) [2] to learn the universal language representations. The $\alpha_{i\tau}$ of the tweet is computed as:

$$h'_{i\tau} = \tanh\left(h_{i\tau}\right) \tag{3}$$

$$\alpha_{i\tau} = \frac{\exp\left(h'_{i\tau} h'^{T}_{W}\right)}{\sum_{m=1}^{M} \exp\left(h'_{im} h'^{T}_{W}\right)} \tag{4}$$

where $h'_{i\tau}$, h'^{T}_{W} are tweet words new hidden state and the learnable parameters, respectively. Similarly, the $e_{l\omega}$ of the retweet is computed as:

$$h'_{l\omega} = \tanh\left(h_{l\omega}\right) \tag{5}$$

$$e_{l\omega} = \frac{\exp\left(h'_{l\omega} h'^{T}_{W}\right)}{\sum_{z=1}^{Z} \exp\left(h'_{lz} h'^{T}_{W}\right)} \tag{6}$$

where $h'_{l\omega}$, h'^{T}_{W} are retweet words new hidden state and learnable parameters, respectively. $\alpha_{i\tau}$ and $e_{l\omega}$ are used to create semantic vectors that reflect the input of tweets and retweets by computing a weighted sum of hidden representations.

User Encoding. As the content of tweets and retweets posted by a user is rather reflected by the user's behavior, the purpose of user encoding is to improve the capturing of users' behavior and personality characteristics as reflected in their profiles.

Recent efforts have revealed that fake news behavior has a strong connection with user popularity [8]. Consequently, we crawl the user characteristics associated with tweets and retweets, including the user's number of followers, friends, and tweets, then model user history using the previously mentioned context modeling technique.

3.2 Temporal Graph Learning Component

Learning a heterogeneous temporal graph can help address two fundamental network interaction representation issues. The first challenge is that the temporal graph based on a tweet and retweet might be sparse since most tweets are directed at the owner of the news article, and retweets are directed at the tweets associated with it. The second challenge is related to the distinct characteristics of fake news - content is repeatedly shared by different entities within a relatively short time. Such challenges may be effectively addressed by simultaneously modeling networks' heterogeneous content and temporal characteristics.

Therefore, the temporal network interaction learning module in Fig. 2 aims to learn the temporal characteristics of interactions between entities and capture the repeating patterns of fake news behavior (Fig. 1). The module achieves this goal by first constructing a temporal interaction graph (Next section). We, then, presents the details of extracting temporal characteristics of various interactions.

Temporal Graph Construction. As shown in temporal graph learning in Fig. 2, we design a temporal network where nodes are news articles, users, tweets, and retweets, and edges are time intervals between the nodes. In particular, each news article has several tweets, and a tweet consists of multiple retweets posted by different users. Such flow is essential for identifying informative connections between nodes in the graph and capturing their temporal characteristics. Thus, we assume the release time of a tweet occurs after the publication date of the news article, and the release time of a retweet occurs after writing date of the tweet.

Table 1. Characteristics of fake and real news

From	To	Edge Weight
Tweets	News articles	Time since article publication
Tweet reply	Original tweet	Time since the original tweet
Retweets	Original tweet	Time since the original tweet
Users	Tweets or retweets	Constant (1)

Since the paper's primary goal is to use the influence of such information flow for fake news detection, we have eliminated tweets and retweets with invalid time and dates to prevent noise information created during feature extraction. Table 1 shows the relationship pairings for the tweet, retweet, and user nodes in the heterogeneous network we construct.

Temporal Engagement Behaviour. The temporal proximity of prescribed interactions in Table 1 is a critical aspect in capturing recurrence. The propagation of fake news via tweets and a fast rupture of retweets can be identified by capturing the time duration among each part of article-tweet and tweet-retweet interactions.

We define a simple but effective weight function (embedded in an adjacency matrix as described in the next section) on the graph's edges to overcome the challenges. Let m denote the time that has been converted to Unix time. ($00:00:00$ UTC on 1 $January 1970$). Given a graph $G = (V, E)$, where V, and

E donate the number of nodes and edges. if the two nodes are connected, the weight q on the edge between them is computed as follows:

$$q_s = m_v - m_{v+1} \tag{7}$$

where s is the encoded edge by the time difference of nanoseconds of both nodes v and $v+1$, where $v \in V$

3.3 Heterogeneous Social Network (HSN)

The ultimate objective of this paper is to enhance fake news detection using heterogeneous interaction models and their temporal engagement interaction. In this section, we detail how to aggregate the tweets, retweets, user, and temporal characteristics of heterogeneous interaction in a coherent framework called the heterogeneous social network. The HSN learns the graph with the proper meta path and then optimizes the fake detection performance.

Inspired by GTN, heterogeneous social network is defined as $G = (V, E)$, where V, and E are the number of nodes and edges, along with node type mapping function $\phi : V \to Q$ and the edge type mapping function $\psi : E \to J$, where Q and J denote the node and relation types, $|Q| + |J| > 2$. Specifically, $Q = a, t, r, u$, where a, t, r, u donate the type of article, tweet, retweet and user, respectively. $J = \{t$ write about $a,\ t$ is shared by $r,\ u$ is related to t, u isrelated to $r\}$. The input to our social network architecture is a set of multiple heterogeneous graph structures, where each graph represents a single news article a along with its associated tweets, retweets, and users. The summary input of one graph structure is as follows:

- **Node features.** We construct our feature matrix $X \in^{\mathbb{D} \times N}$, where N is the concatenated representation of nodes' features, including tweets t_i, retweets r_l, and users characteristics u for each news article. $\mathbb{D}$ represents the combined length of node features. The length of the node feature vector is $\mathbb{D} = l + 3$, where l represents the hidden state dimension of the social context encoder and the additional three dimensions are reserved for user features.
- **Adjacency matrices.** A set of adjacency matrices $\{A\}_{b=1}^{B}$ can represent the heterogeneous graph , where $B = |J|$. $A_b \in^{N \times N}$ is an adjacency matrix where $A_b[m_v, m_{v+1}]$ is non-zero when there is a b^{th} type edge from v to $v+1$. The edge weight represents the temporal features among interactions in the heterogeneous graph as presented in Sect. 3.2.

The main idea of the HSN is to learn a new adjacency matrix A of proper and informative meta-path p linked by specific relation types. This provides an opportunity to identify graph structures with more relevant meta-paths, which can subsequently be utilized to do the classification task.

Based on the above observations, each g^{th} layer in HSN starts to choose adjacency matrices (relation types) by performing 1×1 convolution with the softmax weights function as (part 3 in Fig. 2):

$$F(A;\theta^g) = con_{1\times 1}(A; softmax(\theta^g)) \tag{8}$$

The meta-path adjacency matrix is then produced by multiplying the output matrix with the output matrix of the previous $(g-1)^{th}$ GT Layer.

$$A^{(g)} = (D^{(g)})^{-1} A^{(g-1)} F(A;\theta^{(g)}) \tag{9}$$

where $D^{(g)}$ is a degree matrix that represents the result of multiplying the two matrices $A^{(g-1)}F(A;\theta^{(g)})$. GNNs are deployed for each channel of the output tensor $A^{(g)}$ after the stack of g^{th} layer and then node representation C is updated following the same procedures in [14]. Finally, the label of the fake news y' is computed via several fully connected layers and a softmax layer:

$$y' = Softmax(\mathbf{W}(C) + b) \tag{10}$$

where $\mathbf{W}$ is the learnable parameter and b is the bias value. $y' = [0, 1]$, where 1 is for fake news and 0 for real news.

4 Experiments

This section begins with an overview of the used datasets for detecting false news as described in 4.1. The experimental setup, including baselines, evaluation metrics, and implementations, is then detailed in Sect. 4.2. The predictive performance of T-FND in comparison to existing fake news detection models is detailed in Sect. 4.3, where Sect. 4.4 evaluates the effectiveness of the architecture components on the overall performance.

Table 2. The statistics of PolotiFact and GossipCop dataset

Platform	PolitiFact	GossipCop
# Users	31,524	94,647
# Fake news	346	4,592
# Real news	308	13,925
# Tweets	328,954	106,640,5
# Retweets	707,98	228,360

4.1 Dataset

We use FakeNewsNet [10], a well-known fake news dataset. Obtaining the dataset demands a crawl using these references described by the Kai script [10]. However, we can not crawl the dataset directly due to API changes by Twitter.

Thus, we adjust the script by (i) Updating the old libraries and performing the needed changes and (ii) modifying the code to match Twitter policy. Also, as Twitter limits the number of requests that users are allowed to use the API in an arbitrary time, the total download of the dataset takes about three months. Furthermore, we eliminate the missing news articles and tweets (not available any more). We also find sources of noise. For example, some of the referenced tweets are part of multiple news articles of both classes (real/fake). In this case, we drop these tweets from our dataset. The detailed statistics of the dataset are shown in Table 2

4.2 Experimental Setup

For each dataset, we randomly select 75% of the dataset for training and 25% for testing. The sentence length is 280 characters, while the word length is set to 20 characters. The embedding dimension is set to 200 for textual input, and the dropout rate is 0.2. Each experiment is conducted five times, and the average performance is reported.

Evaluation Metrics. We utilize four well-known evaluation metrics in binary classification tasks: accuracy, precision, recall, and F1 score. These measures have been demonstrated to perform well on imbalanced datasets [9].

Baselines. We compare our proposed model with state-of-the-art models, including the following categories: graph neural network methods (GAT [12], HetGNN [15], and GSAGE [4]), Text classification methods (HAN [13], text-CNN [5]), and fake news detection models (TGNF [11], GCAL [6], and BiGCN [1])

Table 3. Performance Comparison

Method	**PolitiFact**				**GossipCop**			
	Accuracy%	Precision%	Recall%	F1%	Accuracy%	Precision%	Recall%	F1%
GAT	80.2	82.3	79.9	80	78.8	77.4	75.4	74
HetGNN	82.7	77.1	73.2	74.5	80.1	75.5	71.3	72.5
GSAGE	72.5	68.2	70.3	70	71.3	67.3	70.2	69.8
HAN	73.3	74.3	65	68.4	70.4	72.4	62.7	65
Text-CNN	77.8	72.2	74	75.5	69.6	69.1	68.5	67.2
TGNF	90.4	88.2	**91.4**	88	82.6	81.4	80.1	79.9
GCAL	89.35	88.8	86.3	87	81.7	80.7	80.4	81.4
BiGCN	89.1	87.6	86.3	84	80.9	80.1	79.9	80.2
T-FND	**92.5**	**91.2**	91.3	**90**	**84.6**	**83.5**	**82**	**83.1**

4.3 Performance Evaluation

This section describes the overall performance of our proposed T-FND model for detecting fake news. All findings are determined by the average value, which is obtained by repeating the same procedure five times. The detailed evaluation measure values for all techniques are shown in Table 3, along with their respective findings.

Comparing T-FND with other fake news baselines that also consider various kinds of features, such as temporal features, we observe that T-FND outperforms other methods in most circumstances with both datasets. Such predictive success may be driven by the fact that our technique efficiently represents diverse interactions in a graph neural network with an emphasis on the temporal characteristics of different nodes. The generalized representation of our proposed model that can capture temporal characteristics also improve the identification of the repetitive characteristics of fake news behavior. The enhancement further comes from modeling user personality features while learning user interaction.

By evaluating different approaches' performances, we also observe that: (1) GNNs approaches, such as HetGNN, provide considerably better performances than text classification approaches, such as Text-CNN and Text-RNN. (2) Another key is that models that are designed particularly for fake news detection, such as GCAL and BiGCN can further improve the performance. (3) Introducing temporal features improves feature-based models. For instance, TGNF introduces a temporal feature that captures dynamic evolution patterns of news propagation. As a result, the TGNF outperforms GCAL.

Table 4. Ablation study results on Politifact and Gossipcop datasets.

Method	**PolitiFact**				**GossipCop**			
	Accuracy%	Precision%	Recall%	F1%	Accuracy%	Precision%	Recall%	F1%
T-FND w/o TemF	89	88.1	88.2	88.2	82.1	81.2	81	82
T-FND w/o RetF	88.3	87.6	88.4	87.5	82.3	81.7	82.2	82
T-FND w/o UserF	90.2	90.1	90	89.9	83.7	83.4	83.5	82.5
T-FND	92.5	91.8	91.4	90.2	84.6	83.5	82.5	82.2

4.4 Ablation Study

In this section, we conduct experiments to determine the influence of the major components on T-FND. In particular, we compare T-FND to the following variations by deleting specific components from the T-FND model during the model training step: (i) *T-FND w/o TemF*: The temporal dynamic features are removed, (ii) *T-FND w/o RetF*: The retweet context features are removed, and (iii) *T-FND w/o UserF*: We eliminated user features, such as number of followers. The leave-one-out technique is used to examine the influence of each of the three components - retweet content, temporal characteristics, and user information - on T-FND performance, as summarized in Table 4.

Table 4 show that all three aspects are significant for detecting fake news in both datasets. When retweets' content characteristics are eliminated from the learning process, T-FND's performance deteriorates the most. As a result, modeling the interplay between tweets and retweets is critical. T-FND enforces correlations between semantically relevant tweets and retweets by explicitly connecting the tweet content's edge weights to the retweet content's modeling, hence decreasing context-induced noise. For example, introducing the temporal dynamic features improve the performance from 89% to 92.5%. Capturing and representing temporal characteristics (Table 1) in the temporal graph further enhances performance substantially. Moreover, the consideration of user information enhances the performance of T-FND, indicating the effectiveness of modeling users' personalities' behavior. By integrating these three aspects, T-FND can reach the best overall performance.

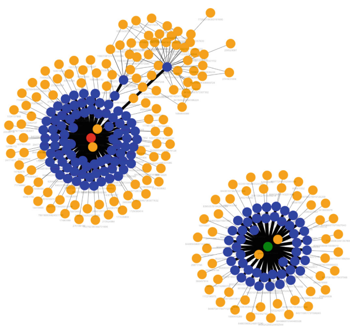

Fig. 3. An example of a propagation graph for two news articles: One for fake news (red node) and the other one for real news (green node). The blue and yellow nodes represent the tweets and retweets, respectively. We can observe the number of retweets for the fake news article is larger than the number of retweets for the real news article. (Color figure online)

5 Discussion

Graphs are usually represented by adjacent matrices. As these adjacent matrices have a shape of N × N (N representing the count of nodes), they become significantly larger when there are many nodes. Detecting the spread of fake news on social media can mean that there are tens of thousands of participants for a single news article. Even if one would only utilize the tweets, the FakeNewsNet dataset has data points with more than 20,000 tweets, resulting in an input shape for the model of more than 20,000 * 20,000. With a footprint of 32 bits, the adjacent matrix itself would have a memory footprint of 1.49 GB. As most

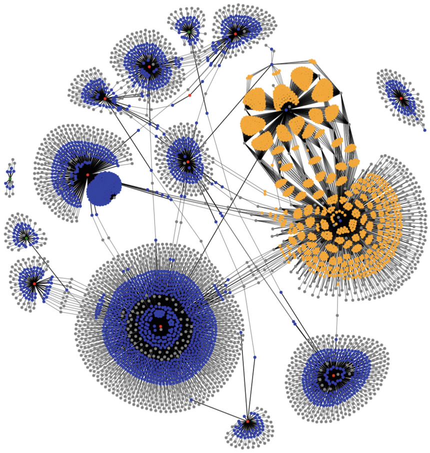

Fig. 4. An example of a propagation graph for random news articles. Red and blue nodes represent the fake and real news articles, respectively. The blue and yellow nodes represent the tweets and retweets, respectively, where the gray nodes refer to the users. We can observe the number of retweets for the fake news article is larger than the number of retweets for the real news article. (Color figure online)

values in the adjacent matrix are zero, one possible solution is to use advanced sparse data structures to ameliorate the situation.

Similar issues with model memory are given when using a large language model to encode the contents of the tweets, as they also have high embedding dimensions per token. With 40 tokens per tweet, 20,000 tweets, 768 embedding dimensions per tweet, and a float 32-bit encoding, the input data of this single example datapoint becomes 2.29 GB. The size of the input data only delivers additional challenges, as model size and gradients have to fit into the memory of GPUs as well. This makes research for small labs only utilizing consumer graphics cards impossible. No optimization technique allowed us to train this model on hardware with 16GB memory. Instead, the training required us to use GPUs with bigger memory.

6 Conclusion

In this paper, we show that modeling tweets, retweets, and temporal interaction is crucial for capturing fake news behavior's repetitive and variation characteristics, resulting in better prediction performance. This paper proposes T-FND, fake news detection framework with a focus on heterogeneous interactions. We empirically evaluate the effectiveness of our model and its components on real-world datasets which outperform state-of-the-art methods for fake news detection.

Our paper opens several key avenues for future work. One such direction is investigating the district sharing pattern by social media platforms, as the number of retweets for fake news is double the number of retweets for real news (Fig. 3 and Fig. 4). Additionally, we can incorporate other important aspects of fake news into the temporal interaction graph. For instance, we can consider different roles of users (spreaders of fake news such as botnets) and study how a user's role evolves over time in the graph modeling process.

Acknowledgment. This work was supported in part by the National Science Foundation Program under award No. 1939725.

References

1. Bian, T., et al.: Rumor detection on social media with bi-directional graph convolutional networks. In: Proceedings of the AAAI Conference on Artificial Intelligence (2020)
2. Dong, L., et al.: Unified language model pre-training for natural language understanding and generation. In: NIPS (2019)
3. Dong, Y., Chawla, N.V., Swami, A.: metapath2vec: scalable representation learning for heterogeneous networks. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (2017)
4. Hamilton, W., Ying, Z., Leskovec, J.: Inductive representation learning on large graphs. In: Advances in Neural Information Processing Systems (2017)
5. Kim, Y.: Convolutional neural networks for sentence classification. In: EMNLP, pp. 1746–1751 (2014)
6. Liao, H., Liu, Q., Shu, K., et al.: Fake news detection through graph comment advanced learning. arXiv preprint arXiv:2011.01579 (2020)
7. Ravazzi, C., Tempo, R., Dabbene, F.: Learning influence structure in sparse social networks. IEEE Transactions on Control of Network Systems (2017)
8. Sharma, K., Qian, F., Jiang, H., Ruchansky, N., Zhang, M., Liu, Y.: Combating fake news: A survey on identification and mitigation techniques. ACM Transactions on Intelligent Systems and Technology (TIST) (2019)
9. Shu, K., Cui, L., Wang, S., Lee, D., Liu, H.: dEFEND: explainable fake news detection. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (2019)
10. Shu, K., Mahudeswaran, D., Wang, S., Lee, D., Liu, H.: FakeNewsNet: a data repository with news content, social context, and spatiotemporal information for studying fake news on social media. Big data (2020)
11. Song, C., Shu, K., Wu, B.: Temporally evolving graph neural network for fake news detection. Inf. Process. Manag. **58**, 102712 (2021)
12. Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., Bengio, Y.: Graph attention networks. In: International Conference on Learning Representations (2018)
13. Yang, Z., Yang, D., Dyer, C., He, X., Smola, A., Hovy, E.: Hierarchical attention networks for document classification. In: Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (2016)
14. Yun, S., Jeong, M., Kim, R., Kang, J., Kim, H.J.: Graph transformer networks. In: NIPS (2019)
15. Zhang, C., Song, D., Huang, C., Swami, A., Chawla, N.V.: Heterogeneous graph neural network. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (2019)