

HG-SL: Jointly Learning of Global and Local User Spreading Behavior for Fake News Early Detection

Ling Sun¹, Yuan Rao^{1*}, Yuqian Lan¹, Bingcan Xia¹, Yangyang Li²

¹ Xi'an Key Laboratory of Social Intelligence and Complexity Data Processing, School of Software Engineering, Xi'an Jiaotong University, China

²National Engineering Laboratory for Risk Perception and Prevention, China

{sunling, Yuqian_Lan_xjtu, bingcan92}@stu.xjtu.edu.cn, raoyuan@mail.xjtu.edu.cn, liyangyang@cetc.com.cn

Abstract

Recently, fake news forgery technology has become more and more sophisticated, and even the profiles of participants may be faked, which challenges the robustness and effectiveness of traditional detection methods involving text or user identity. Most propagation-only approaches mainly rely on neural networks to learn the diffusion pattern of individual news, but this is insufficient to describe the differences in news spread ability, and also ignores the valuable global connections of news and users, limiting the performance of detection. Therefore, we propose a joint learning model named **HG-SL**, which is blind to news content and user identity, but capable of catching the differences between true and fake news in the early stages of propagation through global and local user spreading behavior. Specifically, we innovatively design a **H**ypergraph-based **G**lobal interaction learning module to capture the global preferences of users from their co-spreading behaviors, and introduce node centrality encoding to complement user influence in hypergraph learning. Moreover, the designed **S**elf-attention-based **L**ocal context learning module first introduce spread status in behavior learning process to highlight the propagation ability of news and users, thus providing additional signals for verifying news authenticity. Experiments on real-world datasets indicate that our HG-SL, which solely relies on user behavior, outperforms SOTA baselines utilizing multidimensional features in both fake news detection and early detection task.

Introduction

Serious cases of spreading fake news (Grinberg et al. 2019; Vosoughi, Roy, and Aral 2018) have posed a significant threat to social stability and even national security, aggravated the urgency of developing efficient detection methods.

Classical fake news detection approaches are mainly based on the semantics or styles of news content (Vlachos and Riedel 2014; Wu et al. 2020a). However, given that fake news is deliberately fabricated to mislead consumers, such methods are difficult to identify well-disguised fake news. Therefore, recent researches tried to complement the content features with relevant comments (Shu et al. 2019; Wu et al. 2020c), participants' attributes (Monti et al. 2019), social networks (Min et al. 2022) and propagation structures (Shu

et al. 2020b; Ma, Gao, and Wong 2018). While such strategies enhanced the effectiveness of detection, they will inevitably be fooled by the glorified identities of social robots or the instructive fake comments given by malicious manipulators (Figure 1). (Allcott and Gentzkow 2017).

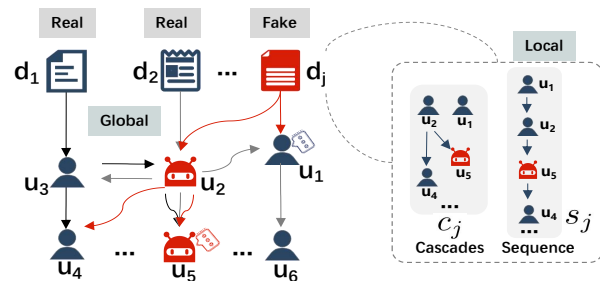


Figure 1: Illustration of global and local user spreading behavior, fake objects and abnormal users are highlighted in red. The abnormal user u_2 tricks a normal user u_4 to spread the fake news d_j , and u_5 cooperates with u_2 to make fake comments, further causing confusion.

Compared with methods involving text or user attributes, deceiving a spreading behavior-only model requires disturbing the holistic propagation patterns involving many ordinary users, which makes such methods theoretically robust (Rosenfeld, Szanto, and Parkes 2020). Recently, a series of models that only rely on propagation have been proposed, such as PPC (Liu and Wu 2018), Pattern-driven approach (Zhou and Zafarani 2019) and WL graph kernel (Rosenfeld, Szanto, and Parkes 2020). However, most of them are proved to be not competitive in detection. We summarize the following possible reasons:

First, few models learn the connections of news and users from a global perspective, which limits the learning of user identities and preferences. Generally, there are differences in the behavior patterns of ordinary users and special accounts (e.g., bots) (Orabi et al. 2020), and given that users' attributes can be faked, it is reasonable and more stable to learn their identities through behavior. In addition, learning the global behavior of users also helps to reveal their preferences on content (individuals tend to believe information that confirm their existing cognitions) (Allcott and Gentzkow 2017) and environment (credibility of the infor-

*Corresponding author.

mation to the individual will increase if others, especially the trust-worthy one tend to believe it) (Dmj et al. 2018), thus providing supplementary clues for fake news detection.

Second, propagation-based methods often rely on neural networks for feature learning (i.e. RNNs for sequence learning and GNNs for structure learning). However, neural networks pay more attention to feature transformation and aggregation, but incapable of capturing the status of propagation itself, such as speed and breadth that reflect the propagation ability of news. While Vosoughi et al. (Vosoughi, Roy, and Aral 2018) proved that spread status do reflect the difference between true and fake news, they demonstrated little the effectiveness of these properties for fake news detection.

To solve the above issues, we propose a novel model that jointly learns the user spreading behavior at global and local levels, and introduces complementary encodings to enhance the learning ability of neural networks, thus obtaining more discriminative representations of true and fake news. Specifically, instead of referring to the previous strategies based on heterogeneous graphs (Yuan et al. 2019) or weighted union graphs (Tu et al. 2021), we introduce **Hypergraph** to describe users' **Global interactions (HG)**. Since each hyperedge can link an arbitrary number of entities, hypergraph helps to simultaneously learn users' preferences. In **Self-attention-based Local context learning module (SL)**, we highlight the local context under a specific news through multi-head self-attention mechanism, and with the spread status encoding, our model can simultaneously gain insight into the propagation ability of users and news during the learning process. The main contributions of our work are as follows:

- We propose a novel fake news detection model that jointly learns the global and local user spreading behavior through Hypegraph Neural Network and Multi-head Self-attention, and demonstrate that the propagation patterns without texts and user profiles can provide powerful signals for revealing news veracity.
- We are the first to introduce the spread status of news in neural network training process, which highlight the propagation ability of news and users, and further enhance the descriptive ability of neural networks.
- Experimental results on two real-world datasets show that HG-SL significantly outperforms previous state-of-the-art detection methods. Moreover, as a robust model using fewer features, HG-SL enables efficient and stable detection in the early stage of propagation.

Related Work

The fake news detection task aims to distinguish whether a news spread on online social platforms is fake based on relevant information, such as news content, users' comments, participants' identities, propagation patterns, etc.

Early approaches focus on linguistic differences. Besides shallow features (Kakol, Nielek, and Wierzbicki 2017), news semantics (Wu et al. 2020a), style (Gröndahl and Asokan 2019), and sentiment (Giachanou, Rosso, and Crestani 2019) have also been explored. However, fake news is designed to mimic the real one, and as the forgery technology improves, the detection effect of such methods becomes

weak. Therefore, other relevant texts have been considered in recent studies. For instance, dDEFEND (Shu et al. 2019) and DTCA (Wu et al. 2020c) develop sentence-comment network to exploit semantic conflicts of news contents and user comments; EHIAN (Wu et al. 2020b) tries to find evidence from news and relevant articles. These strategies enhance the interpretability and stability of detection, but may still suffer from interference from bot accounts (Gilani et al. 2019) and "alternative media" (malicious websites that frequently release false or highly biased posts (Starbird 2017)). In addition, users' biases against content and other users also inevitably bring noise to such methods as they may lead users to express inaccurate opinions.

In view of the differences in the spread of fake and real news, propagation features have been utilized to enhance detection (Shu et al. 2020b; Zhou and Zafarani 2019; Liu and Wu 2018; Bian et al. 2020). Vosoughi et al. (Vosoughi, Roy, and Aral 2018) analyzed the spread of news reports on Twitter and found that falsehood diffused significantly farther, faster, deeper, and broader than the truth in all fields, even though the statistical indicators they summarized proved to not perform well on the detection task (Rosenfeld, Szanto, and Parkes 2020), their findings attracted researchers' attention to the propagation mode of true and fake news. Zhou et al. (Zhou and Zafarani 2019) then concluded four patterns to reflect the nature of fake news, namely More-Spreader, Further-Distance, Stronger-Engagement and Denser-Network. Rather than relying on handcrafted features, PPC (Liu and Wu 2018) models the propagation as multivariate time series, and only relies on the attributes of participants. Instead of using RNNs, Bian et al. (Bian et al. 2020) designed a Bi-GCN model to model the bidirectional propagation trees. GCNFN (Monti et al. 2019) utilizes users' profiles to supplement the comment embeddings, and UPFD (Dou et al. 2021) further captures the historical posts of users to represent their endogenous preferences. The effectiveness of this type of approach is undeniable, but its high demands on data cannot be ignored.

Compared to methods that consider news content or the identities of associated users, models based on propagation alone appear to be more stable, as they are less likely to be cheated by fictitious texts or identities. Rosenfeld et al. (Rosenfeld, Szanto, and Parkes 2020) designed a Weisfeiler-Lehman graph kernel that is blind to text, user and time, and proved that topologically encodings of cascades provide rich clues for predicting news credibility. However, since such model relies on fewer features, its detection performance still lags behind models using multi-dimensional features.

Problem Formulation

We let $D = \{d_1, d_2, \dots, d_m\}$ to represent the news set, m is the total number of news. The collection of users participating in news propagation is denoted as $U = \{u_1, u_2, \dots, u_n\}$. For global learning, we construct a hypergraph $G = (U, E)$ to describe users' global interactions at the news level, E represents the set of hyperedges. Each hyperedge e_j connects all users that tweet or retweet the j -th news d_j . For local learning, we define the propagation cascades and sequences of news D as $C = \{c_1, c_2, \dots, c_m\}$ and $S =$

$\{s_1, s_2, \dots, s_m\}$, separately. $c_j = \{c_{j,1}, c_{j,2}, \dots, c_{j,k}\}$ represents the propagation cascades of d_j , which contains one or more propagation trees $c_{j,p} = \{(u_i, L_i^{j,p}, I_i^{j,p}) | u_i \in U\}$, we use $L_i^{j,p}$ and $I_i^{j,p}$ to denote the depth and the number of child nodes of user u_i in $c_{j,p}$, thus preserving the structural features. $s_j = \{(u_i, t_i^j) | u_i \in U\}$ represents the propagation sequence of news d_j , t_i^j indicates the timestamp of u_i spreading d_j . Each news d_j is assigned with a label $y_j \in \{0, 1\}$, if news d_j is fake, $y_j = 1$, otherwise $y_j = 0$. The task of our work is to predict the label of d_j by learning the hypergraph G , cascades c_j and sequence s_j .

The Proposed Model

The overall architecture of HG-SL is shown in Figure 2. Instead of directly using user attributes for preferences learning (Monti et al. 2019; Liu and Wu 2018), which may be disturbed by fabricated identities, we construct global propagation hypergraph to capture users' preferences more robustly from their behavior patterns. With the addition of node centrality encoding, the global influence of users will be highlighted. Since the local propagation context cannot be obtained by graph learning, multi-head self-attention modules with spread status encoding are designed to learn local news propagation from structural and temporal aspects respectively, then the two embeddings are combined by gated fusion for a more comprehensive expression.

Hypergraph-based Global Interaction Learning

The bot-like signs and preferences of users can be reflected from their behaviors and connections, which imply the credibility of users and provide valuable clues for fake news detection. Therefore, we construct a hypergraph to describe the global co-spreading behavior of users, and utilize Hyper-GNN and node centrality encoding for hypergraph learning.

Node Centrality Encoding Graph models always emphasize the transformation and aggregation of node attributes, resulting in the loss of the structural characteristics of nodes. As a strong signal to measure the global importance of users in network, node centrality is introduced to enhance the learning ability of neural network. Since centrality indicators on simple graphs such as degree centrality and closeness centrality do not apply to hypergraphs, we define the activity degree as the centrality of user in hypergraph since active users provide richer information:

Activity degree: the total number of hyperedges that the node u_i participates in: $Act_i = |\mathcal{E}_i|$, where \mathcal{E}_i is the set of hyperedges containing node u_i .

To incorporate the centralities into the training process of Hyper-GNN, we use an embedding function to generate centrality vector Cen_i from the activity degrees, which will be directly added to the original embedding to obtain $x_i^0 = x_i^{ini} + Cen_i$, the initial embedding x_i^{ini} is randomly initialized from normal distribution.

Hypergraph Neural Network (Hyper-GNN) We use a hypergraph neural network with two-stage aggregation to

model the global behavior of users. Note that the hyperedge itself does not contain any features, it is only used to assist node aggregation, i.e. we do not learn the content of news.

Nodes-to-edge Aggregation. For each hyperedge e_j , the first step of Hyper-GNN aims to learn its representation a_j by aggregating the embeddings of all its connected nodes:

$$\mathbf{a}_j^l = \sigma \left(\sum_{u_i \in e_j} \frac{1}{|e_j|} \mathbf{W}_1 \mathbf{x}_i^{l-1} \right) \quad (1)$$

where σ is the activation function **ReLU**, $\mathbf{W}_1 \in \mathbb{R}^{d^d \times d^d}$ is the trainable weight matrix, d^d is the dimension of embedding, l is the layer of Hyper-GNN.

Edges-to-node Aggregation. Then we train another aggregator to integrate all hyperedges \mathcal{E}_i participated by node u_i to update the representation of node u_i :

$$\mathbf{x}_i^l = \sigma \left(\sum_{e_j \in \mathcal{E}_i} \frac{1}{|\mathcal{E}_i|} \mathbf{W}_2 \mathbf{a}_j^l \right) \quad (2)$$

where σ is activation function **ReLU**, $\mathbf{W}_2 \in \mathbb{R}^{d^d \times d^d}$ is trainable weight matrix. After the two-stage aggregation, the updated representation of node u_i contains not only its own information, but also the information of nodes that have shared news with it, which reflects its global preference.

Self-attention-based Local Context Learning

Hyper-GNN focuses on the global relations of news and users, but it is incapable to describe the internal context under a specific news. Therefore, we integrate the spread status into two multi-head self-attention modules to learn the local representation of news from structural and temporal aspects.

Local Temporal Learning The details of temporal learning are illustrated in Figure 3. Temporal encodings of users and sequence are introduced as complement before and after self-attention learning, respectively.

Temporal encoding of users. We preserve the timestamp t_i^j of each user u_i participating in the sequence s_j to reflect the time differences between participants, and utilize an embedding function to generate vector $tu_{i,1}^j$ for the timestamp. Since the timestamps are not continuous, we use the absolute order of participation as position information for the training of self-attention, and encode it as $tu_{i,2}^j$. The above two embeddings will be directly added to x_i to obtain temporal-aware representation of news d_j : $\mathbf{o}_j^{T'} = [(x_i + tu_{i,1}^j + tu_{i,2}^j) | u_i \in s_j]$.

Multi-head Self-Attention Based on the outstanding performance of self-attention mechanism in sequential tasks, we apply multi-head self-attention module to learn the local context of propagation. The basic learning process is:

$$\text{Att}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax} \left(\frac{\mathbf{Q}\mathbf{K}''}{\sqrt{d^d/H}} \right) \mathbf{V} \quad (3)$$

where H denotes the number of attention heads, \mathbf{K}'' is transpose of \mathbf{K} . The learned representation h_j^T is calculated as:

$$\begin{aligned} \mathbf{h}_{q,j}^T &= \text{Att} \left(\mathbf{o}_j^{T'} \mathbf{W}_q^{QT}, \mathbf{o}_j^{T'} \mathbf{W}_q^{KT}, \mathbf{o}_j^{T'} \mathbf{W}_q^{VT} \right) \\ \mathbf{h}_j^T &= [\mathbf{h}_{1,j}^T; \mathbf{h}_{2,j}^T; \dots; \mathbf{h}_{H,j}^T] \mathbf{W}_O^T \end{aligned} \quad (4)$$

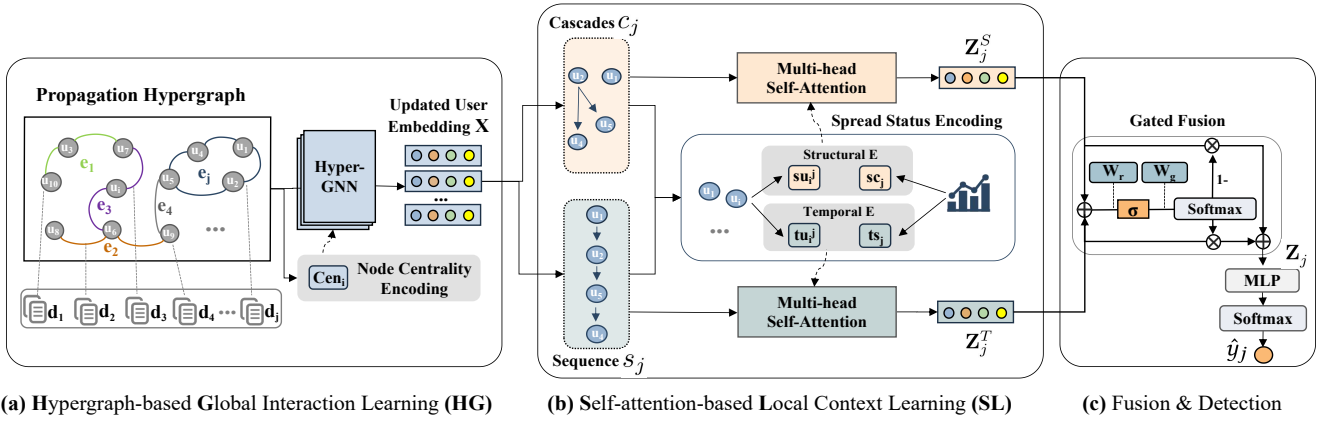


Figure 2: An overview of the architecture of HG-SL which consists of three major components: (1) Global interaction learning module uses hypergraph neural networks and node centrality encoding to learn the global relations of users, (2) local structural and temporal features are learned in local context learning module through multi-head self-attention mechanism and spread status encoding, and (3) in fusion & detection module, news propagation representations from structural and temporal aspects are merged for detection through gated fusion mechanism.

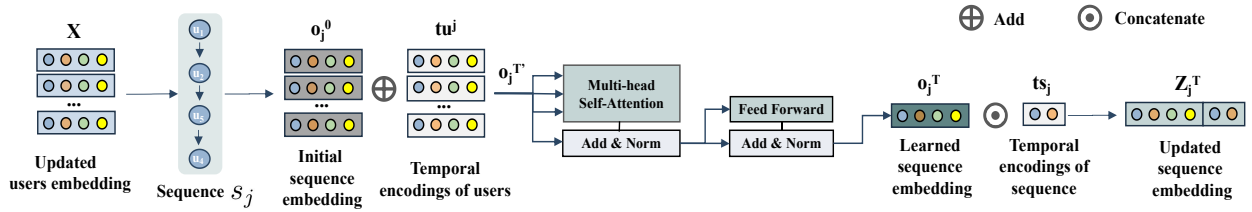


Figure 3: Temporal learning process of multi-head self-attention module. User-level temporal encodings are introduced before learning, while encodings of sequence level are supplemented after self-attention learning.

where \mathbf{W}_q^{QT} , \mathbf{W}_q^{KT} , \mathbf{W}_g^{VT} and \mathbf{W}_O^T are learnable matrices. Then we use a feed forward network (two layers fully-connected neural network) to obtain the learned sequence embedding, and take the mean value as the finally \mathbf{o}_j^T :

$$\mathbf{o}_j^T = \text{MEAN}(\mathbf{W}_{A_2} \sigma(\mathbf{W}_{A_1} (\mathbf{h}_j^T)'' + \mathbf{b}_1) + \mathbf{b}_2) \quad (5)$$

in which σ is the activation function ReLU , \mathbf{W}_{A_1} and \mathbf{W}_{A_2} are learnable matrices, \mathbf{b}_1 and \mathbf{b}_2 are bias parameters.

Temporal encoding of sequence. Since the duration of spread ($ts_{j,1}$) and the average response time from tweet to retweet ($ts_{j,1}$) help to reflect the propagation speed of news d_j , we take the above two features as sequence-level temporal features. Given that the above features are of numerical float type, we directly concatenate them as complementary features to the sequence representations learned by self-attention (\mathbf{o}_j^T). Finally, the time-aware representation of news d_j is denoted as $\mathbf{Z}_j^T = [\mathbf{o}_j^T, ts_j] \in \mathbb{R}^{d^{d'}}$, $d^{d'} = d^d + 2$ is the updated dimension.

Local Structural Learning Similar to temporal learning, we train another multi-head self-attention module with structural encodings to obtain structure-aware local news representation.

Structural encoding of users. The number of retweets caused by user u_i in sub cascade $c_{j,p}$ is indicated as struc-

tural feature at the user level to highlight the local importance of u_i . Moreover, the depth of u_i in $c_{j,p}$ will be provided as position information to the self-attentional learning process. We use two embedding functions to generate structural embeddings $su_{i,1}^{j,p}$ and $su_{i,2}^{j,p}$ from the user importance and position, respectively. They will be directly added to the embeddings x_i to get $\mathbf{o}_j^{S'} = [(\mathbf{x}_i + su_{i,1}^{j,p} + su_{i,2}^{j,p}) | u_i \in c_{j,p}]$.

Structural encoding of cascades. Given that a news d_j may generate multiple cascades in propagation, we use the number of sub-cascades in c_j ($sc_{j,1}$), and the proportion of non-isolated cascades $sc_{j,2}$ to represent the breadth and attractiveness of news propagation, concatenate them with the news cascades representation learned by the multi-head self-attention module (\mathbf{o}_j^S), and finally obtain $\mathbf{Z}_j^S = [\mathbf{o}_j^S, sc_j]$.

Fusion & Detection

Gated Fusion To incorporate the learned structural and temporal local propagation features for a more expressive representation, we introduce a gated fusion mechanism which adaptively combine the two representations as:

$$\mathbf{Z}_j = g\mathbf{Z}_j^S + (1 - g)\mathbf{Z}_j^T \quad (6)$$

$$g = \frac{\exp(\mathbf{W}_g \sigma(\mathbf{W}_r \mathbf{Z}_j^S))}{\exp(\mathbf{W}_g \sigma(\mathbf{W}_r \mathbf{Z}_j^S)) + \exp(\mathbf{W}_g \sigma(\mathbf{W}_r \mathbf{Z}_j^T))}$$

Datasets	PolitiFact	GossipCop
# Fake news	157	2,732
# Real news	157	2,732
# Users	28,049	65,064
# Spreads	36,481	291,043
Avg. # Participants / News	116	53
Max # Participants / News	759	195
Min # Participants / News	1	1

Table 1: Statistics of datasets used in our experiments

where \mathbf{W}_g and \mathbf{W}_r are the transformation matrix and vector for attention respectively, σ is the activation function \tanh .

Fake News Detection Finally, the Softmax function is used to calculate the probability that news d_j is fake:

$$\hat{y}_j = \text{softmax}(\mathbf{W}_p \mathbf{Z}_j + \mathbf{b}_p) \quad (7)$$

where \mathbf{W}_p is the transformation matrix, and \mathbf{b}_p denotes the bias. Training data with real labels are used to minimize the cross entropy loss:

$$\mathcal{J}(\theta) = -\frac{1}{m} \left(\sum_{j=1}^m y_j \log(\hat{y}_j) + (1 - y_j) \log(1 - \hat{y}_j) \right) \quad (8)$$

in which θ represents all parameters that need to be learned. $y_j = 1$ means news d_j is fake, otherwise $y_j = 0$.

Experiments

To validate the effectiveness of our proposed HG-SL model, we conduct extensive experiments on real datasets to answer the following research questions:

- **RQ1.** How does the proposed HG-SL perform on the fake news detection task compared to previous works?
- **RQ2.** Can HG-SL identify fake news at the early stage of propagation?
- **RQ3.** What are the contributions of jointly learning and other components to the performance of HG-SL?

Datasets

Following the previous works (Shu et al. 2020b; Dou et al. 2021), we utilize the public fake news detection data repository FakeNewsNet (Shu et al. 2020a), which consists of news data related to two fact-checking websites: GossipCop and PolitiFact. News in PolitiFact mainly involves political topics, while GossipCop mainly includes entertainment news. The original datasets include rich information such as the text of news, user tweet, retweet and comment behavior, and the timestamps of users’ engagements. Given that our model aims to find differences in propagation patterns between true and fake news, we only choose the structural and temporal information of users’ tweets and retweets behavior for feature learning (i.e. positions and timestamps). The statistics of the sampled datasets are listed in Table 1.

Baselines

We compare HG-SL with the following detection methods:

- **GRU**(Ma et al. 2016): a RNN-based model that learns temporal patterns from propagation sequence.
- **PPC**(Liu and Wu 2018): uses recurrent and convolutional networks to learn users’ attributes and propagation paths to detect fake news.
- **CSI**(Ruchansky, Seo, and Liu 2017): employs LSTM to encode the news content, and utilizes the group behavior of users who propagate fake news for detection.
- **BiGCN**(Bian et al. 2020): leverages a top-down and a bottom-up GCN to learn the patterns of rumor propagation and rumor dispersion, respectively.
- **GCNFN**(Monti et al. 2019): encodes the directed news propagation graph through extended GCN, and takes the comment and profile information as the user feature.
- **GLAN**(Yuan et al. 2019): models the relationships among source tweets, retweets, and users as a heterogeneous graph to capture the rich structural information.
- **UPFD**(Dou et al. 2021): learns user preferences through their past engaged posts, and combines content with graph modeling.
- **HPFN**(Shu et al. 2020b): classifies news through machine learning classifiers and the extracted propagation features from structural, temporal, and linguistic aspects.

Evaluation Metrics and Parameter Settings

As a binary classification task, the accuracy (Acc.), precision (Prec.), recall (Rec.) and F1 score are adopted for evaluation. Our experiments are conducted on a GPU device (12 GB GeForce GTX 2080Ti). For each dataset, we randomly choose 20% of the news for training, 10% are used for validation, and the remaining 70% are held to evaluate the model performance in the test phase. The maximum number of user engagements is set to be 200. Due to the data limitation, for GRU, CSI, BiGCN, GCNFN and GLAN, we only leverage user profiles provided in (Dou et al. 2021) as user features. In terms of UPFD, we utilize the embeddings of users’ past engaged posts as their features. We preserve the model settings as provided in original papers. For our proposed HG-SL, we implement it in PyTorch and adopt Adam as the optimizer. For the Politifact dataset, we need to train for 200 epochs, while the Gossipcop dataset only needs 20 epochs to achieve optimal performance. The learning rate is set to be 0.001, the dropout rate is 0.5 and the batch size is 64. The dimension of user and news embedding d is set to be 64. We tried {1, 2, 3}-layer Hyper-GNN to learn the global hypergraph and found that a single layer Hyper-GNN is sufficient to capture the critical connections with the two-round aggregation strategy. The number of heads H used in multi-head self-attention is set to be 8, which adjusted in {2, 4, 8, 12, 16}. We ignore the analysis of hyperparameters due to page limitations.

Performance Comparison (RQ1)

We compare HG-SL with the baselines on two public datasets, the results are shown in Table 2. Specifically, we can observe that: (1) Our model achieves an accuracy of 90.05% on Politifact dataset and 98.04% on Gossipcop dataset, respectively, which are 5.7% and 0.95%

Method	Considerations					Politifact				Gossipcop				
	Text	Temp.	Stru.	User	Local	Global	Acc.	Prec.	Rec.	F1	Acc.	Prec.	Rec.	F1
GRU	✓	✓			✓		62.73	63.82	61.30	60.98	78.92	82.33	78.57	80.36
PPC		✓		✓	✓		64.21	64.56	61.07	62.82	90.38	90.33	91.10	90.71
CSI	✓	✓		✓	✓	✓	75.33	83.36	74.20	75.82	78.20	82.34	78.32	80.89
BiGCN			✓	✓	✓		78.83	80.39	76.26	78.71	89.02	92.79	84.81	88.97
GCNFN	✓		✓	✓	✓		82.35	86.85	76.80	82.24	95.61	94.40	97.02	95.59
GLAN	✓		✓	✓	✓	✓	82.34	84.68	83.59	84.13	95.52	94.63	93.29	93.96
UPFD	✓		✓	✓	✓	✓	84.31	87.14	81.03	84.25	97.09	96.91	97.31	97.07
HPFN	✓	✓	✓		✓		75.63	71.46	85.40	77.80	86.39	85.66	87.41	86.53
Ours		✓	✓		✓	✓	90.05	92.30	88.61	89.93	98.04	98.40	97.68	98.01

Table 2: Performance comparison of our proposed HG-SL with baselines (%).

higher than the second best model UPFD. The outstanding results indicate that our strategy to jointly learns the global and local user spreading behavior can effectively capture the difference between true and fake news. (2) Of the three approaches that focus on temporal feature while ignoring propagation structures (GRU, PPC and CSI), CSI analyzes both content and user behavior, and thus performs better on Politifact dataset with insufficient data. Moreover, the PPC model relying on user features outperforms the content-based GRU model on both datasets, proving the importance of user behavior information in detection. (3) GLAN, BiGCN, GCNFN and UPFD are propagation structure-based approaches, among them, GLAN combines local text features and global propagation features, so it outperforms BiGCN that only learns local propagation trees. In addition, GCNFN and UPFD optimize the text embedding and graph learning strategy and thus achieved slightly better results than GLAN. UPFD additionally introduces the user’s historical preference as a global endogenous feature, thus achieving best performance among baselines (84.31% and 97.09% respectively). (4) The feature engineering-based method HPFN calculates the propagation features from content, temporal, and structural aspects, and achieves the accuracy of 75.63% and 86.39% on two datasets, even better than some deep learning models, which strongly demonstrates the effectiveness of spread status information for detection.

Early Detection (RQ2)

Early detection aims to identify fake news as early as possible, thereby minimizing the impact of fake news. We define two early detection scenarios: limiting the number of user engagements (tweets /retweets) and limiting the detection deadline, and carry out comparative experiments to further prove the effectiveness and stability of our model.

The impact of limitation of user engagements Referring to Fig. 4, we observe that the detection accuracy of all models on two datasets increases steadily with the addition of engagements, and HG-SL always gets the highest score. It is worth noting that our model can reach accuracies of 77.37% and 93.28% on Politifact and Gossipcop datasets with only the first 10 engagements, which already exceeds the best performance achieved by GRU, PPC and CSI using 200 engagements, demonstrating the effectiveness of combining the global and local propagation features.

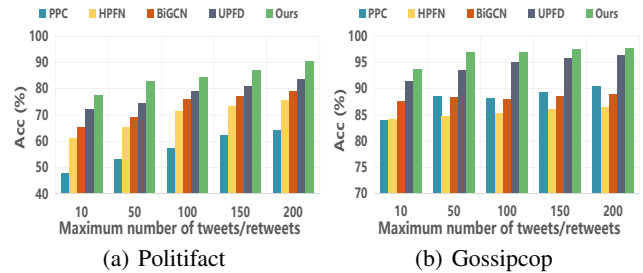


Figure 4: Performance comparison under different maximum engagements (tweets/retweets).

The impact of detection deadline Unlike early detection scenarios that limit the number of user engagements, limiting detection deadlines allows faster-spreading news to contain more training data. The relationships between the detection deadline and the average number of user engagements on the two datasets are shown in the Fig. 5(a) and 5(c). Note that although the overall average number of engagements in Politifact (116) is much higher than that of Gossipcop (53), the news it contains spreads more slowly, with an average of only 9.49 engagements per news in the first 4 hours, and only 38.18 engagements at 36 hours, compared to 22 and 41.18 on Gossipcop. This predicts that the detection on Gossipcop will achieve stability faster. Referring to Fig.5(b) and 5(d), HG-SL using less than 4-hour data (79.85%) to outperform the best baseline UPFD using data in 24 hours (78.53%) on Politifact. In terms of Gossipcop, HG-SL taking only the first 4 hours of data (95.27%) even exceeds HPFD’s detection performance at 36 hours (95.21%), indicating the effective early detection performance of our model.

Ablation Study (RQ3)

We conduct ablation studies over the different parts of HG-SL to investigate the contribution of submodules, the results are reported in Table 4. The variants are designed as:

- **HG** ignores global learning and removes Hyper-GNN and global centrality encoding .
- **SL** ignores local learning and removes self-attention modules and spread status encoding.
- **Structural SL** ignores local structural learning.
- **Temporal SL** ignores local temporal learning.
- **Node centrality E** removes global centrality encoding.

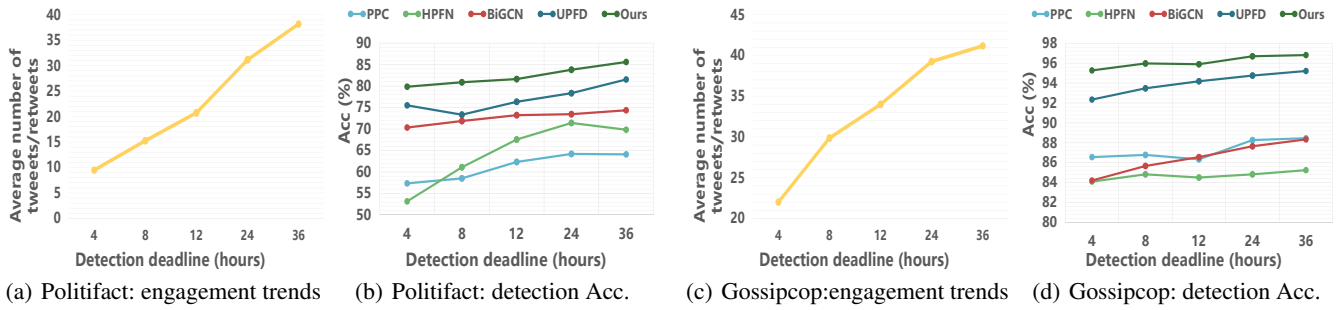


Figure 5: Performance comparison with different detection deadlines

Encodings		Politifact							Gossipcop						
		Fake			Real			P-value	Fake			Real			P-value
		Min	Max	Avg	Min	Max	Avg		Min	Max	Avg	Min	Max	Avg	
Stru.	SC ₁	1	282	79.4	1	450	62.7	0.05	1	173	26.1	1	187	60.6	0.0
	SC ₂	0	1	0.17	0	1	0.19	0.30	0	1	0.15	0	1	0.04	8e-198
Temp.	TS ₁	0	3517.6	262.5	0	4143.5	1495.1	3.9e-21	0	3692.6	322.3	0	3555.3	62.7	1.5e-72
	TS ₂	0	1195.1	32.5	0	3355.4	490.2	4.1e-12	0	3692.5	137.9	0	3239.6	21.9	2.9e-34

Table 3: Analysis of statistical spread status features. (p-value less than 0.05 is significant)

Method	Politifact		Gossipcop	
	Acc.	F1	Acc.	F1
- HG	78.73	78.07	96.00	95.97
- SL	85.52	85.43	97.69	97.67
- Structural SL	88.23	88.11	97.80	97.78
- Temporal SL	85.97	85.91	97.83	97.80
- Node centrality E	89.59	89.55	97.72	97.70
- Structural E	88.68	88.65	97.88	97.86
- Temporal E	87.97	87.84	97.85	97.83
- Gated fusion	89.59	89.46	97.97	97.58
HG-SL	90.05	89.93	98.04	98.01

Table 4: Ablation study (%).

- **Structural E** removes local structural encoding.
- **Temporal E** removes local temporal encoding.
- **Gated fusion** replaces gated fusion with addition.

As shown in Table 4, HG-SL generally achieves the best performance compared to any of its variants, indicating the rationality of its design. Specifically, the results prove that: **Effectiveness of joint learning:** The removal of either global learning or local learning significantly degrades the performance of HG-SL, which demonstrates the rationality of our joint learning strategy. Specifically, the model shows the biggest drop in performance after removing the Hyper-GNN, with a drop of around 11.3% on Politifact and 2% on Gossipcop, which proves that users’ global interactions do help to characterize their preferences. Moreover, the self-attention mechanism emphasize local context within the propagation, therefore, the removal of it also has a great impact on results, especially on Politifact, which shows about a 4.5% drop in both accuracy and F1 scores. **Effectiveness of encodings:** Removing any encoding degrades the model’s performance, which demonstrates that introducing spread status of news and users indeed enhances the learning ability

of neural networks. Furthermore, we conduct an analysis on the selected statistical news spread status in Table 3 to intuitively explain why they are useful. The analysis reveals that the differences in the spread of true and fake news persist on different platforms (check the p-values), but the differences are not always fixed across platforms. To illustrate, Politifact is consistent with the findings of Vosoughi et al., that is, fake news always spreads faster, farther, deeper, and wider. It is reflected in our statistics that fake news has shorter spread period (TS₁), faster spread speed (TS₂), contains more cascades (SC₁) and less independent cascades (SC₂) in propagation. However, we observe opposite phenomenon on Gossipcop, which implies why the statistical features are not very efficient as decisive indicators for detection, but help to supplement the behavior learning in our model. **Effectiveness of gated fusion:** Unlike the addition operation which simply merges two vectors, our gating function controls the retaining rate of two vectors through correlation calculation, thus improving the performance on both datasets.

Conclusion and Future Work

In this work, we propose a novel joint learning model for fake news detection task named HG-SL. To improve the reliability of detection and go beyond the limitations of previous methods based on unilateral propagation features, we use Hyper-GNN to embed users’ global relations, and meanwhile utilize multi-head self-attention modules to learn the local context within a propagation, so as to comprehensively capture the difference between true and fake news. The introduced global node centrality and local spread status further highlight the influence of users and the spread ability of news. Experiments show that HG-SL can significantly outperform SOTA models on fake news (early) detection task.

In the future, we plan to consider other behaviors and the stances of users to improve the interpretability of detection.

Acknowledgments

The research work is supported by National Key Research and Development Program in China (2019YFB2102300), and the National Natural Science Foundation of China (Grants No. U22B2036).

References

- Allcott, H.; and Gentzkow, M. 2017. Social Media and Fake News in the 2016 Election. *Nber Working Papers*, 31(2): 211–236.
- Bian, T.; Xiao, X.; Xu, T.; Zhao, P.; Huang, W.; Rong, Y.; and Huang, J. 2020. Rumor Detection on Social Media with Bi-Directional Graph Convolutional Networks. In *Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020*, 549–556.
- Dmj, L.; MA, B.; Y, B.; AJ, B.; KM, G.; F, M.; MJ, M.; B, N.; G, P.; and D, R. 2018. The science of fake news. *Science*, 359(6380): 1094–1096.
- Dou, Y.; Shu, K.; Xia, C.; Yu, P. S.; and Sun, L. 2021. User Preference-aware Fake News Detection. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '21*, 2051–2055.
- Giachanou, A.; Rosso, P.; and Crestani, F. 2019. Leveraging Emotional Signals for Credibility Detection. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2019, Paris, France, July 21–25, 2019*, 877–880. ACM.
- Gilani, Z.; Farahbakhsh, R.; Tyson, G.; and Crowcroft, J. 2019. A Large-scale Behavioural Analysis of Bots and Humans on Twitter. *ACM Transactions on the Web*, 13: 1–23.
- Grinberg, N.; Joseph, K.; Friedland, L.; Swire, B.; and Lazer, D. 2019. Fake news on Twitter during the 2016 U.S. presidential election. *Science*, 363: 374–378.
- Gröndahl, T.; and Asokan, N. 2019. Text Analysis in Adversarial Settings: Does Deception Leave a Stylistic Trace? *ACM Comput. Surv.*, 52(3): 45:1–45:36.
- Kakol, M.; Nielek, R.; and Wierzbicki, A. 2017. Understanding and predicting Web content credibility using the Content Credibility Corpus. *Inf. Process. Manag.*, 53(5): 1043–1061.
- Liu, Y.; and Wu, Y. B. 2018. Early Detection of Fake News on Social Media Through Propagation Path Classification with Recurrent and Convolutional Networks. In *Proceedings of the 32th AAAI Conference on Artificial Intelligence, (AAAI-18)*, 354–361.
- Ma, J.; Gao, W.; Mitra, P.; Kwon, S.; Jansen, B. J.; Wong, K.; and Cha, M. 2016. Detecting Rumors from Microblogs with Recurrent Neural Networks. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI 2016, New York, NY, USA, 9–15 July 2016*, 3818–3824. IJCAI/AAAI Press.
- Ma, J.; Gao, W.; and Wong, K. 2018. Rumor Detection on Twitter with Tree-structured Recursive Neural Networks. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, ACL 2018*, volume 1, 1980–1989.
- Min, E.; Rong, Y.; Bian, Y.; Xu, T.; Zhao, P.; Huang, J.; and Ananiadou, S. 2022. Divide-and-Conquer: Post-User Interaction Network for Fake News Detection on Social Media. In *WWW '22: The ACM Web Conference 2022, Virtual Event, Lyon, France, April 25 - 29, 2022*, 1148–1158. ACM.
- Monti, F.; Frasca, F.; Eynard, D.; Mannion, D.; and Bronstein, M. M. 2019. Fake News Detection on Social Media using Geometric Deep Learning. *CoRR*, abs/1902.06673.
- Orabi, M.; Mouheb, D.; Aghbari, Z. A.; and Kamel, I. 2020. Detection of Bots in Social Media: A Systematic Review. *Inf. Process. Manag.*, 57(4): 102250.
- Rosenfeld, N.; Szanto, A.; and Parkes, D. C. 2020. A Kernel of Truth: Determining Rumor Veracity on Twitter by Diffusion Pattern Alone. In *Proceeding of the Web Conference 2020, WWW 2020*, 1018–1028.
- Ruchansky, N.; Seo, S.; and 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, CIKM 2017, Singapore, November 06 - 10, 2017*, 797–806. ACM.
- Shu, K.; Cui, L.; Wang, S.; Lee, D.; and Liu, H. 2019. dE-FEND: Explainable Fake News Detection. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2019*.
- Shu, K.; Mahudeswaran, D.; Wang, S.; Lee, D.; and Liu, H. 2020a. 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.
- Shu, K.; Mahudeswaran, D.; Wang, S.; and Liu, H. 2020b. Hierarchical Propagation Networks for Fake News Detection: Investigation and Exploitation. In *Proceedings of the Fourteenth International AAAI Conference on Web and Social Media, ICWSM 2020*, 626–637.
- Starbird, K. 2017. Examining the Alternative Media Ecosystem Through the Production of Alternative Narratives of Mass Shooting Events on Twitter. In *Proceedings of the Eleventh International Conference on Web and Social Media, ICWSM 2017*, 230–239.
- Tu, K.; Chen, C.; Hou, C.; Yuan, J.; Li, J.; and Yuan, X. 2021. Rumor2vec: A rumor detection framework with joint text and propagation structure representation learning. *Inf. Sci.*, 560: 137–151.
- Vlachos, A.; and Riedel, S. 2014. Fact Checking: Task definition and dataset construction. In *Proceedings of the Workshop on Language Technologies and Computational Social Science, ACL 2014*, 18–22.
- Vosoughi, S.; Roy, D.; and Aral, S. 2018. The spread of true and false news online. *Science*, 359(6380): 1146–1151.
- Wu, L.; Rao, Y.; Nazir, A.; and Jin, H. 2020a. Discovering differential features: Adversarial learning for information credibility evaluation. *Inf. Sci.*, 516: 453–473.
- Wu, L.; Rao, Y.; Yang, X.; Wang, W.; and Nazir, A. 2020b. Evidence-Aware Hierarchical Interactive Attention Networks for Explainable Claim Verification. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI 2020*.

Wu, L.; Rao, Y.; Zhao, Y.; Liang, H.; and Nazir, A. 2020c. DTCA: Decision Tree-based Co-Attention Networks for Explainable Claim Verification. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020*.

Yuan, C.; Ma, Q.; Zhou, W.; Han, J.; and Hu, S. 2019. Jointly Embedding the Local and Global Relations of Heterogeneous Graph for Rumor Detection. In *Proceedings of the IEEE International Conference on Data Mining, ICDM 2019*, 796–805. IEEE.

Zhou, X.; and Zafarani, R. 2019. Network-based Fake News Detection: A Pattern-driven Approach. *SIGKDD Explor.*, 21(2): 48–60.