

Research Article

Heterogeneous Graph Convolutional Network-Based Dynamic Rumor Detection on Social Media

Dingguo Yu ^{1,2}, Yijie Zhou ^{1,3}, Suiyu Zhang ², and Chang Liu ^{1,3}

¹Key Lab of Film and TV Media Technology of Zhejiang Province, Communication University of Zhejiang, Hangzhou 310018, China

²College of Media Engineering, Communication University of Zhejiang, Hangzhou 310018, China

³Institute of Intelligent Media Technology, Communication University of Zhejiang, Hangzhou 310018, China

Correspondence should be addressed to Dingguo Yu; yudg@cuz.edu.cn

Received 10 January 2022; Revised 21 February 2022; Accepted 19 March 2022; Published 26 April 2022

Academic Editor: Keke Shang

Copyright © 2022 Dingguo Yu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The development of social media has provided open and convenient platforms for people to express their opinions, which leads to rumors being circulated. Therefore, detecting rumors from massive information becomes particularly essential. Previous methods for rumor detection focused on mining features from content and propagation patterns but neglected the dynamic features with joint content and propagation pattern. In this paper, we propose a novel heterogeneous GCN-based method for dynamic rumor detection (HDGCN), mainly composed of a joint content and propagation module and an ODE-based dynamic module. The joint content and propagation module constructs a content-propagation heterogeneous graph to obtain rumor representations by mining and discovering the interaction between post content and propagation structures in the rumor propagation process. The ODE-based dynamic module leverages a GCN integrated with an ordinary differential system to explore dynamic features of heterogeneous graphs. To evaluate the performance of our proposed HDGCN model, we have conducted extensive experiments on two real-world datasets from Twitter. The results of our proposed model have outperformed the mainstream model.

1. Introduction

Recent years have witnessed the prevalence of mobile Internet technology and social networks, which provide a convenient platform for users to obtain a large amount of information and express their opinions. Since more people have gotten involved in hot topics and exchanged their views on social media, the rumors have been circulated. Meanwhile, rumors can be manipulated to spread widely and mislead users for certain purposes [1, 2], which can damage the reputations of individuals and have a detrimental effect on society [3–5]. For example, during the COVID-19 outbreak, the world has been facing a large number of health-related rumors about coronavirus. One of the typical rumors is that a traditional Chinese medicine called Shuanghuanlian can prevent novel coronavirus, causing many citizens to queue to buy overnight [6]. The presence of such rumors on social media seriously affected social stability. Therefore,

it is urgent to come up with automatic detection methods to identify rumors accurately and timely.

Most of the early detection methods adopt handcrafted feature engineering to extract features from user profiles [7, 8], content [9–12], and propagation patterns [13–15] to train supervised classifiers. Corresponding methods are used for training, such as Decision Tree [16], Random Forest [8], and Support Vector Machine [7]. However, these methods that rely on manual feature engineering consume both time and workforce. Moreover, those methods cannot capture the higher-dimensional feature representations from rumor propagation.

Recently, deep learning has been successfully applied in computer vision, natural language processing, and recommender system. Given its ability to automatically extract low- and high-order feature representations, many studies have exploited deep learning methods in rumor detection. The deep learning model, long short-term

memory (LSTM), and gated recurrent unit (GRU) [17] are applied to capture sequential features from rumor propagation along time. It is the first time introducing deep learning to obtain temporal representations of rumor propagation. Then, Ma et al. [18] designed a tree-based recursive neural network (RvNN) to learn the semantic sequence and propagation structure from two directions for rumor detection. These approaches focus on sequential propagation but neglect dispersion. Bruna et al. [19] pointed out that Convolutional Neural Network (CNN) could obtain the relevant features of local neighbors. Therefore, CNN-based methods [20, 21] were put forward to mine the dispersion structure of the rumor propagation, obtaining local correlation features representation from neighbor retweets. However, these methods cannot deal with the global structure of rumor propagation. In order to capture global structure feature representations during propagation, graph convolutional network (GCN) is proposed [22]. Since GCN is widely used in complex networks such as social networks and physical systems, Wu et al. [15] proposed an undirected GCN model to obtain global structural features of rumor propagation. Still, they do not take the direction of propagation into account. Therefore, Bian et al. [23] leveraged a GCN with a bidirectional graph network to learn the global features of patterns and diffusion of rumor propagation. These methods mentioned above only take propagation structure into consideration and ignore rumor content features which are also vital for rumor detection [24, 25]. Since these methods above do not take the heterogeneous information into consideration, a heterogeneity network-based method [26] is proposed to tackle the complex network, which makes positive performance to link prediction. Huang et al. [27] proposed a heterogeneous method by constructing a tweet-word-user heterogeneous graph to make the most of content and propagation information. Besides, rumor propagation is a continuous-time dynamic system [28]. Recent studies on rumor detection do not pay attention to dynamic changes during rumor propagation. Nevertheless, it is essential to delve into dynamic changes of rumor propagation structure for better performance. Shang et al. [29] found limitations of the static graph in analyzing social network and explored interaction evolution of the dynamic network. To date, some studies have focused on using ordinary differential equations to capture the continuous-time dynamic of the complex networks [30–32] in other fields, such as fluid mechanics [33, 34]. Yildiz [35] proposed a method based on ordinary differential equations and an autoencoder to learn high-dimensional complex continuous-time latent dynamics. Choi et al. [36] leveraged neural ordinary differential equations (NODEs) to learn user and product coevolving embedding over time. The NODEs consider the time variable t as continuous, which enables us to obtain the continuous-time dynamic [37]. The results of the above NODEs demonstrate better performance than only the neural network method. This is because NODEs combine the high-capacity function approximation of deep learning and the high-efficiency dynamic acquisition of differential

equations, which offers the best-of-both-worlds approach to obtain dynamic features. The typical architecture of NODEs is shown in Figure 1.

To deal with both heterogeneous information and dynamic features of rumor detection, in this paper, we propose a novel heterogeneous graph convolutional network-based method for dynamic rumor detection (HDGCN), which integrates an ordinary differential equation system with GCNs to capture the dynamic evolution of the heterogeneous graph. We first construct a heterogeneous graph based on combining content and propagation structures with learning rumor representations, as shown in Figure 2. Then, we design an ODE-based graph convolutional network to learn dynamic representations of nodes during the propagation process. Finally, we introduce a meaning-pooling layer and a fully connected layer to predict rumor classes.

Compared with existing methods, the proposed approach HDGCN is capable of (1) exploiting rumor representation with joint content and propagation pattern; (2) precisely modeling continuous-time dynamic relevance by an ordinary differential equation system to obtain the instantaneous rate of changes in the heterogeneous graph; (3) effectively capturing and leveraging propagation trending and content for rumor classification. To summarize, the main contributions of this work are as follows:

- (1) This novel method constructs the joint post content and rumor propagation structure as a heterogeneous graph to obtain rumor representation, containing textual information and initial propagation information.
- (2) We explore a novel ODE-based graph convolutional network method that integrates an ordinary differential system and GCNs to acquire dynamic changes of the heterogeneous graph. Moreover, this is the first work to leverage the differential equation to obtain the dynamic evolution of heterogeneous graph for rumor detection.
- (3) We perform extensive experiments on real-world Twitter15 and Twitter16 datasets. The result demonstrates better effectiveness of our proposed method HDGCN compared to mainstream methods.

The rest of the paper is organized as follows: in Section 2, related works are reviewed; in Section 3, details of our proposed model HDGCN are introduced; we present the experimental results and discussions in Section 4; finally, Section 5 is the conclusion.

2. Related Work

In recent years, automatic rumor detection on social media has attracted considerable attention. Previous works for rumor detection on social media have focused on hand-crafted feature engineering, extracting rumor features from text content, user profiles, and propagation structures to design a classifier. Castillo et al. [16] pioneered engineering features and classified the rumor using Support Vector Machine and so on, which represented the starting point for

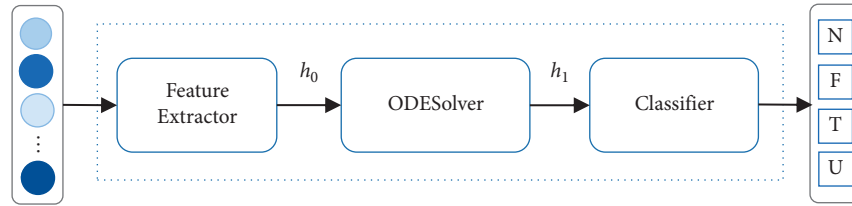


FIGURE 1: The typical architecture of NODEs. Feature extractor provides the initial state h_0 , and h_1 is the dynamic state calculated by NODE methods.

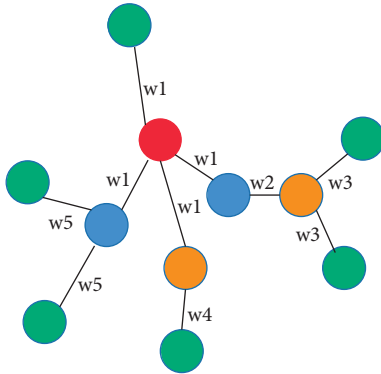


FIGURE 2: The heterogeneous graph combines post content and rumor propagation structure as rumor representation. The edge between two nodes is the content representations built with the word frequency and the inverse post frequency. The red point is the original post. The nodes sharing the same father node have the same rumor content information.

many later works. Yang et al. [7] formulated rumor classification by extending emotional features to enhance the strength of rumor credibility assessment. Dayani et al. [38] proposed a method for obtaining user features and content features by K-Nearest neighbor and Naive Bayes classifier, respectively. Wu et al. [15] modeled the propagation structures by using the random walk method to extract features of the propagation process. Ma et al. [14] constructed a propagation tree kernel to detect rumors by evaluating similarities among the structures of the propagation tree. However, these methods all rely on manual feature engineering to obtain valid information, which is inefficient and cannot extract high-dimensional feature representations.

In order to automatically learn high-dimensional features, many rumor detection methods based on deep learning have been proposed. For propagation structures, Ma et al. [17] leveraged recurrent neural networks (RNN) to capture hidden feature representations of variation tweet content over time. Liu et al. [39] proposed a time-series classifier that incorporates both CNN and RNN, capturing the global and local variations of content features and user features, respectively. Ma et al. [18] designed a tree-based recursive neural network (RvNN) to learn the semantic sequence and propagation structure from two directions for rumor detection. Guo et al. [40] proposed a hierarchical bidirectional long short-term memory model combined with an attention mechanism for learning representations.

For content preprocessing, Varshney et al. [41] developed an automated system to distinguish fake and real news from content resemblance features. Kaliyar et al. [42] proposed the FakeBERT model based on TextCNN [24, 25] and BERT [43] to learn news textual information for detection. Apart from text content, Meel et al. [44] proposed a multimodal fake news detection framework to analyze multimodal data, such as textual information and image features. Besides, Varshney et al. [45] also proposed a scheme to extract text information from the videos to classify fake videos with user profiling and human consensus features.

However, deep learning models mentioned above have no ability to process graph-structural data of rumor propagation, but graph neural network lends itself to learning graph features [22]. Wu et al. [15] proposed a GCN model to obtain global structural features of rumor propagation. Malhotra et al. [46] leverage GCN to classify the constructed propagation graph and use RoBERTa-based word embedding [47] to understand linguistic cues. Additionally, Dou et al. [48] proposed a method based on BERT [43] and GCN, which captures various signals from user preferences. These methods, however, are inefficient in learning the dynamics of rumor propagation.

Compared to the GCN models mentioned above, neural ODEs (NODEs) contribute to capturing dynamic features from neural network architecture. NODEs are able to learn the derivative of hidden state with neural networks and change discrete hidden layers into the continuous-depth network and continuous-time latent variable models for supervised learning [49]. Learning from neural ODEs, Ma et al. [50] leveraged entity recognition, sentence reconfiguration, and an ODE network to construct a rumor detection framework. It is the first time introducing neural ODE into rumor detection field and getting a promising performance. Zang et al. [28] combined an ODE and GNNs to learn continuous-time dynamics on complex networks for predicting continuous-time network dynamics. Yildiz [35] proposed a latent second-order ODE method integrated with an autoencoder to learn high-dimensional trajectories and complex continuous-time latent dynamics. In the recommendation field, Choi et al. [36] extended the linear GCN method with neural ordinary differential equations (NODEs) to learn user and item coevolving embedding over time. Besides, Wang et al. [51] proposed a temporal influence kernel function based on the derivation and decay exponential function to obtain the dynamics of users' preference from sequential records, which derives from the idea of differential equations.

3. Methodology

3.1. Preliminaries. We now introduce a formal description of our problem settings and preliminary knowledge to understand our work. The notations and symbols used are summarized in Table 1.

We define the rumor detection dataset $X = \{x_1, x_2, \dots, x_{|X|}\}$, where each x_i corresponds to the i -th event and $|X|$ is the number of all events. Each event comprises two independent sets, namely, rumor content and propagation structure. So the event x_i can be denoted as $x_i = \{P_i, G_i\}$, where P_i and G_i represent the content and propagation structure of i -th event. The i -th event propagation structure is denoted as $G_i = \{C_i, E_i, A_i\}$, where $C_i = \{c_{i0}, c_{i1}, c_{i2}, \dots, c_{iN}\}$ represents the nodes in the propagation structure of event x_i . c_{i0} is the original post, and c_{ij} is the j -th responsive post during propagation. $E_i = \{e_{st} | s, t = 0, 1, 2, \dots, N\}$ denotes a set of all links. Then, according to the nodes and edges, $A \in \{0, 1\}^{N \times N}$ is denoted as an adjacency matrix. Specifically, $A_{st} = 1$ means the edge from node c_{is} to c_{it} exists. Then, combining content and propagation structure, we set A' as the adjacency matrix of the heterogeneous graph, which is detailed in the next section.

Moreover, we formulate this task as a supervised classification problem. Each event is associated with a ground-truth y_i and learns classifier f from ground-truth labels; that is, $f: x_i \rightarrow y_i$, where y_i is one of the four classes {NR, FR, TR, UR}, representing nonrumor, false rumor, true rumor, and unverified rumor, respectively. Our goal is to train a model $f(\cdot)$ to predict the label of a given post and propagation.

3.2. Model Overview. The basic idea of HDGCN is to learn dynamic changes from joint post content and rumor propagation structure for accurate rumor classification. In this section, we present the construction of our proposed framework for rumor detection based on the continuous-time dynamic on rumor heterogeneous graph.

As shown in Figure 3, the proposed rumor detection model consists of three main components: rumor representation capturing, dynamic propagation modeling, and rumor classification. Firstly, HDGCN encodes rumor representations by constructing a joint content and propagation structure graph. In order to build the heterogeneous graph, we learn rumor content representations via TF-IDF [52], encode propagation node vectors via adjacency matrix, combine content and propagation information together, and then embed the union graph into a low-dimensional space. Secondly, we use an undirected GCN to learn the initial status of the heterogeneous graph. Meanwhile, we devise an ordinary differential equation system with GCNs to obtain continuous-time dynamics of the heterogeneous graph. Finally, a mean-pooling layer is used to concatenate the hidden dynamic features and predict rumor classification. Next, we will elaborate on the details of each component.

3.3. Rumor Representations Capturing. We build a joint rumor content and propagation structure graph, named

TABLE 1: Symbols used in this work.

Symbol	Description
$x_i \in X$	x_i is the i -th event in the event set X as well as the joint graph
P_i	P_i is the i -th post content
G_i	G_i is the propagation graph of the i -th event x_i
$A \in \{0, 1\}^{N \times N}$	The adjacency matrix of rumor propagation graph
A'	The adjacency matrix of heterogeneous graph
A_{st}	The edge weight from node c_{is} to c_{it}
y_i	The label of event x_i
$f(\cdot)$	Mapping functions

content-propagation heterogeneous graph, to encode rumor representation. The heterogeneous graph combines rumor content information (the weights of edges) and rumor propagation structures, which is shown in Figure 2.

To obtain rumor content representation for each post, we adopt the TF-IDF method, which is used for text mining by static word frequency and inverse document frequency. Firstly, we filter the stopped words and construct the corpus. Then, we formulate the term frequency as $tf_{ij} = n_{ij} / \sum_k n_{ik}$, where n_{ij} is the number of times the i -th word appeared in post j . And the inverse document frequency is formulated as $idf_i = \log|D| / 1 + |\{k: t_i \in p_k\}|$, where D denotes the total number of posts in the corpus and $|\{k: t_i \in p_k\}|$ is the number of posts containing word t_i . The weight of word t_i is computed as $t_i = tf_{ij} * idf_i$. So the content representation of post j can be denoted as $p_j = [t_1, t_2, \dots, t_{|W|}]$, where $|W|$ is the number of corpora. Then, because the representation of p_j is high-dimension and sparse, we employ an embedding layer to map them into low-dimension space to acquire dense real-value vectors. Formally, the dense vector of each rumor content is defined as

$$v_j = W_j p_j, \quad (1)$$

where W_j denotes the weight of embedding layer and v_j is considered as the content representation of post j .

Next, we can construct the rumor propagation structure $G_j = \{C_j, E_j, A\}$ based on the spreading path of retweets and responses. Since adjacency matrix $A \in \{0, 1\}^{N \times N}$ can be on behalf of rumor spreading, we can obtain the adjacency matrix A' of joint rumor content and propagation graph from A and v_j , which is defined as

$$A'_{st} = A_{st} v_j, \quad (2)$$

where A_{st} is the value of edge from node c_{js} to c_{jt} .

Then, we can obtain the initial rumor representation r_j , formulated as

$$r_j = \frac{1}{M} \sum_{c_{js}, c_{jt} \in C_j} A'_{st}, \quad (3)$$

where M is the number of edges in the propagation of post j .

3.4. Initial State Acquiring. After obtaining rumor representation, we acquire the initial state of the heterogeneous by

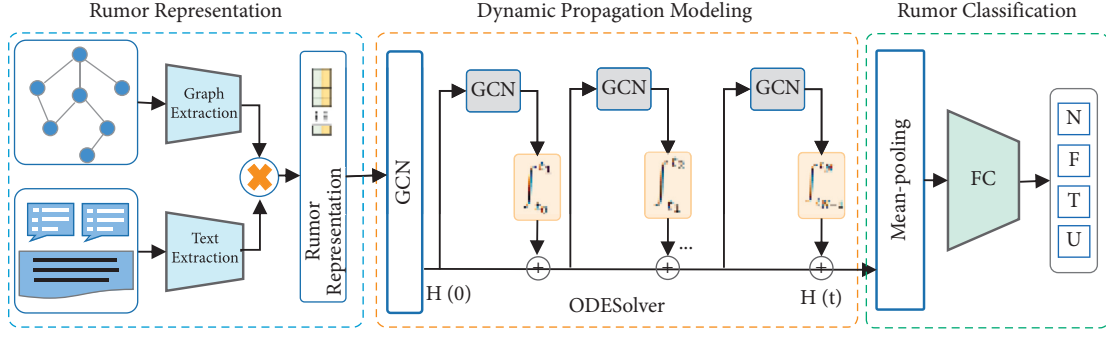


FIGURE 3: The workflow of the proposed model HDGCN. HDGCN consists of three steps: (1) obtaining rumor representation by joint content and rumor propagation structure graph, (2) modeling dynamic propagation by ODE-based GCN, and (3) predicting rumor classification.

a GCN method. We firstly build the operators of convolutional graph neural network \tilde{A} , which is defined as

$$\Phi = \tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}}, \quad (4)$$

$$\tilde{A} = A' + I,$$

where A' and I are the adjacency matrix and identity matrix of the heterogeneous graph, respectively. \tilde{D} is the degree matrix of \tilde{A} .

Then, we obtain the initial state representation of the union graph. The equation is written for the model as follows:

$$H_0 = \sigma(\Phi r_j W_0), \quad (5)$$

where H_0 represents the initial state of rumor propagation. W_0 is the filter weight. And we adopt the ReLU function as the activation function $\sigma(\cdot)$. The dropout is applied in the GCN layer to avoid overfitting.

3.5. ODE-Based Dynamic Propagation Modeling. As we know, the propagation of rumor events can be seen as a continuous-time dynamic complex network. To obtain the dynamic features in joint rumor content and propagation graph, we combine the ordinary differential equation system and graph convolutional networks to learn the continuous-time dependency of propagation progress.

Firstly, we describe the dynamic system of the propagation by a differential equation, formulated as [53]

$$\frac{dH(t)}{dt} = f(H, G, W, t), t > 0, \quad (6)$$

where $H(t)$ represents the state of the propagation network at time t , with $t \in [0, \infty]$. G denotes the heterogeneous graph, capturing how nodes interact with each other. W are parameters which are the weight demonstrating how the progress evolves over time. $dH(t)/dt$ is the derivative of time, denoting the instantaneous rate of rumor propagation dynamic system changes. $H(0) = H_0$ is initial states of propagation progress at time $t=0$.

The nonlinear mapping function $f(H, G, W, t)$, capturing the dynamic changes, is modeled by a GCN, which is defined as

$$f(H, G, W, t) = \sigma(\Phi H_0 W_0 + b), \quad (7)$$

where Φ is the operators of convolutional graph neural networks, W_1 denotes the parameters of GCN, b is the bias, and $\sigma(\cdot)$ is the ReLU activation function.

After getting the initial state, we encode the graph network node state H_t from the original hidden state H_0 and obtain a continuous-time dynamic state at an arbitrary time by integrating $dH(t)/dt$ over time t , which is defined as

$$H_t = H_0 + \int_0^t f(H, G, W, \tau) d\tau, \quad (8)$$

where H_t denotes the current state at arbitrary time t in rumor propagation. The rumor propagation is a time-varying dynamic system in the differential system when coefficient W changes over time. And we can obtain the hidden state at any time by setting time t .

3.6. Rumor Classification. With the rumor dynamic hidden representation, we employ a mean-pooling operator to aggregate information from an ordinary differential equation system grouped by each interval time, which is formulated as

$$H = \frac{1}{T} \sum_{t=0}^T H(t), \quad (9)$$

where T is the number of time points $t \in \{t_1, t_2, \dots, t_T\}$ of the ODE solver.

Then, we apply a fully connected layer and a softmax layer for rumor classification, which is defined as

$$\hat{y} = \text{softmax}(W_{FC} H + b_{FC}), \quad (10)$$

where W_{FC} and b_{FC} are the weight and bias of the last hidden layer. \hat{y} denotes a vector of probabilities for all classes used to predict the label of the rumor event.

We train our model by minimizing the cross-entropy of the predictions \hat{y} and ground-truth distributions y . L_2 regularizer is applied in the loss function to prevent overfitting.

4. Experiments

In this section, we conduct the experiment on two real-world datasets to evaluate the effectiveness of our proposed HDGCN method in comparison with several baseline models.

4.1. Data Description. For experimental evaluation, we use two public datasets collected from real-world social media to verify our proposed method. They are Twitter15 and Twitter16 [14], which contain 1490 and 818 propagation graphs for each and are annotated with four classes. In the graph, nodes represent users, and edges refer to replies and retweets. Each graph is labeled as one of the four classes, that is, nonrumor, true rumor, false rumor, and unverified rumor. The statistic of the two datasets is given in Table 2.

4.2. Baselines. We compare our proposed model HDGCN with some baselines on rumor detection, which is shown as follows:

DTC [16]: this is a method using a Decision Tree classifier based on manually engineering features to obtain the tweet information.

RFC [54]: this is a method using a Random Forest classifier that uses features related to user, linguistic, and structure characteristics of news for rumor detection.

SVM-TS [55]: this is a method based on an SVM classifier with a constructed time-series kernel of hand-crafted features.

SVM-HK [15]: this is a rumor detection method with an SVM classifier using a graph kernel based on the similarity of propagation structures.

GRU-RNN [17]: this is a rumor detection method based on recurrent neural networks with GRU units to capture the variation of contextual information of relevant tweet posts over time.

BU-RvNN and TD-RvNN [14]: a rumor detection method adopts two-directional tree-structured recursive neural models based on a top-down and a bottom-up tree neural network with GRU units to learn rumor information.

Rumor2vec [56]: a rumor detection method adopts a CNN-based model to combine textual content with propagation structures, capturing joint representation learning for rumor detection.

ESODE [50]: a rumor detection method integrates semantic information and an ordinary differential equation network to obtain statistical features for rumor detection.

4.3. Experiment Setup. We implement all deep learning models using PyTorch and GCN models with PyTorch Geometric package. We use TF-IDF [52] with 5000 words for the content representation of the posts. Embedding sizes of nodes and hidden layers are searched in [64,128,256,512]. At last, we choose 128 as our model embedding size, and the batch size is the same size. The categorical cross-entropy is chosen as the loss function. The maximal tolerance of the ODE solver is 0.01, and the time intervals are set to 5

TABLE 2: Statistics of the datasets.

Statistic	Twitter15	Twitter16
Number of posts	331612	204820
Number of original tweets	1490	818
Number of nonrumors	372	205
Number of true rumors	374	203
Number of false rumors	370	205
Number of unverified rumors	374	205

TABLE 3: Experimental results of rumor detection on Twitter15/ Twitter16. (NR: nonrumor; FR: false rumor; TR: true rumor; UR: unverified rumor).

Method	Acc.	NR F_1	FR F_1	TR F_1	UR F_1
(a) Twitter15 dataset					
DTC	0.454	0.733	0.355	0.317	0.415
RFC	0.565	0.810	0.422	0.401	0.543
SVM-TS	0.544	0.796	0.472	0.404	0.483
SVM-HK	0.493	0.650	0.439	0.342	0.336
GRU-RNN	0.641	0.684	0.634	0.688	0.571
BU-RvNN	0.708	0.695	0.728	0.759	0.653
TD-RvNN	0.723	0.682	0.758	0.821	0.654
Rumor2vec	0.796	0.883	0.746	0.836	0.723
ESODE	0.824	0.778	0.834	0.888	0.789
HDGCN	0.834	0.853	0.868	0.859	0.823
(b) Twitter16 dataset					
DTC	0.473	0.254	0.080	0.190	0.482
RFC	0.585	0.752	0.415	0.547	0.563
SVM-TS	0.574	0.755	0.420	0.571	0.526
SVM-HK	0.511	0.648	0.434	0.473	0.451
GRU-RNN	0.633	0.617	0.715	0.577	0.527
BU-RvNN	0.718	0.723	0.712	0.779	0.659
TD-RvNN	0.737	0.662	0.743	0.835	0.708
Rumor2vec	0.852	0.857	0.769	0.927	0.850
ESODE	0.851	0.771	0.856	0.927	0.857
HDGCN	0.865	0.820	0.863	0.930	0.863

snapshots $\{H(t_1), H(t_2), \dots, H(t_5) | 0 < = t_1 < t_2 < \dots < t_5 < = T\}$ in $[0, T]$ of the continuous-time dynamic propagation graph. We use DOPRI5 [57] method as the ODE solver to calculate the differential equation in our HDGCN model. We grid search the best terminal time $T \in [1.5, 2.5]$ and get our terminal time $T=2$. To make a fair comparison, we split the datasets into five parts and conduct 5-fold cross-validation to obtain robust results. For Twitter15 and Twitter16 datasets, we evaluate accuracy over four categories and F_1 on each class. Besides, we optimize the model by the Adam algorithm. The dropout is 0.5, and the learning rate is 0.005. The training is iterated upon 100 epochs. To prevent overfitting, we apply early stopping when the validation stop is reduced by 10 epochs.

4.4. Performance Comparison. In this section, the result of our proposed method HDGCN is compared with the baselines on Twitter15 and Twitter16 datasets. From Table 3, we can draw the following observations.

Firstly, it is shown that the baselines based on deep learning methods like GRU-RNN, BU-RvNN, Rumor2vec,

TABLE 4: The rumor detection performances of with or without dynamic systems on Twitter15 and Twitter16. -wo refers to “without” ODE-based dynamic module.

Method	Acc.	NR F_1	FR F_1	TR F_1	UR F_1
(a) Twitter15 dataset					
HDGCN-wo dynamic	0.785	0.784	0.791	0.796	0.762
HDGCN	0.834	0.783	0.868	0.859	0.823
(b) Twitter16 dataset					
HDGCN-wo dynamic	0.793	0.744	0.772	0.883	0.756
HDGCN	0.865	0.781	0.863	0.930	0.863

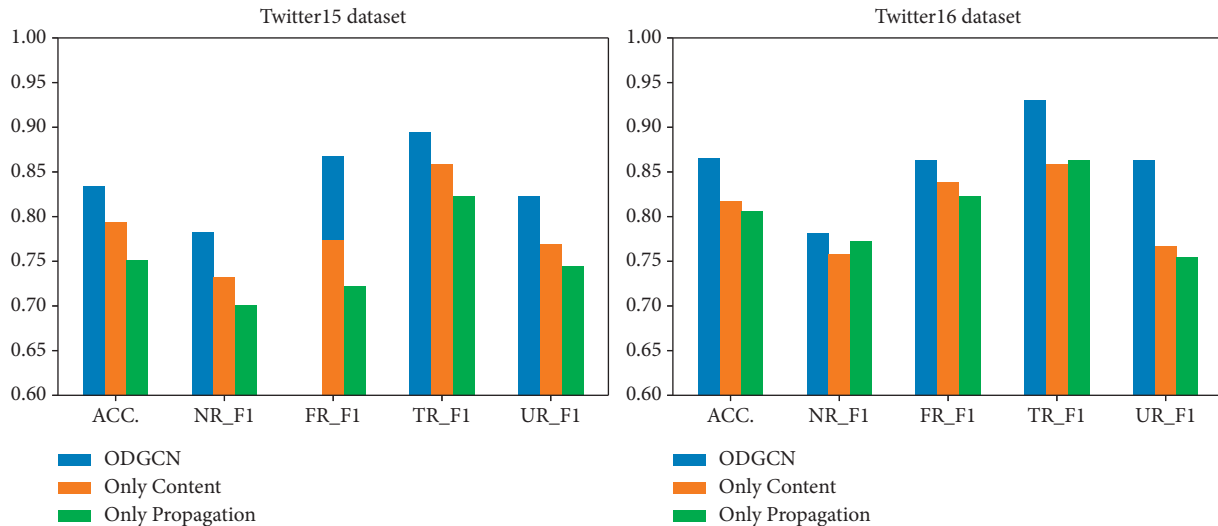


FIGURE 4: The contribution analysis of post content and rumor propagation structures.

and our model outperform the methods using traditional handcrafted features, such as DTC, RFC, SVM-TS, and SVM-HK. This is because the deep learning methods have an excellent ability to automatically learn effective and high-dimensional feature representations of rumors, which demonstrates that the deep learning method benefits from improving the effectiveness of rumor detection.

Secondly, our proposed method HDGCN performs best on accuracy among all baselines on two datasets. Our model achieves an accuracy of 83.4% and 86.5%, respectively, increasing by 4.8% and 1.6% compared with the latest baselines Rumor2vec. The comparison between Rumor2vec and our model demonstrates the effectiveness of incorporating the dynamic features into the model for rumor detection. Since CNN cannot process data with dynamic features, Rumor2vec ignores the importance of dynamic changes during rumor propagation, which causes worse accuracy for rumor detection. Additionally, only for the nonrumor class, the performance of Rumor2vec is better than our proposed model. Because nonrumors are always released or responded by users who have higher authority and credibility, the paths and content of nonrumor propagation are more likely to be fixed. So the dynamics of nonrumor propagation do not significantly affect the performance.

Finally, our method HDGCN is superior to the ESOE method with a 1.2% and 1.6% improvement on two datasets. As we all know, rumor propagation is seen as a graph

structure. The neural ODE method only uses NODE to get dynamic features of propagation but cannot process the graph-structural data, which prevents it from obtaining efficient propagation features, resulting in worse performance for rumor detection.

4.5. Impact of ODE-Dynamic. To analyze the effect of the dynamic features of the heterogeneous graph, we compare the proposed rumor detection method with the variants without the dynamic system. The results are shown in Table 4. By capturing the continuous-time dynamics on the heterogeneous graph, our HDGCN achieves better performance than without dynamic module, increasing by 6.2% and 9.1% on Twitter15 and Twitter16, respectively, which indicates that the dynamic features of the heterogeneous graph play an important role in rumor detection. This is because the GCN model has only learned static representations of graph structures and failed to capture the continuous-time dynamic features precisely. In contrast, the differential equation system can capture the dynamic features by its derivative operator and allow our method to pay more attention to dynamic changes over time, which helps to improve the performance of our model.

Moreover, the effect of dynamic features on the nonrumor class is less than that of other classes, which implies fewer dynamic features of the nonrumor propagation

structure. This is because the nonrumors are usually released or responded by an authority or government department, and the propagation paths and content are more likely to be fixed. Therefore, the dynamics of nonrumor propagation have less effect on the performance.

4.6. Ablation Study. We perform the ablation studies over our heterogeneous graph to analyze the importance of post content and propagation structure for rumor detection. The experimental results are presented in Figure 4. “Only content” refers to our model with post content features in the network, while “only propagation” represents our model with propagation structure only in the heterogeneous graph.

As shown in Figure 4, we can observe that the heterogeneous graph with post content and propagation performs better than the only content or only propagation on accuracy and F_1 of Twitter15 and Twitter16. This indicates that both content and propagation structure benefit from increasing accuracy of rumor classification. Moreover, the performance of most “only content” is better than that of “only propagation.” Since the dynamic change of rumor content may change the rumor meanings, post content dynamics play a more critical role in rumor detection than propagation dynamics. In general, it is necessary to combine the content and propagation structure for rumor detection.

5. Conclusion

In this paper, we propose a graph convolutional network-based dynamic rumor detection method named HDGCN, which takes advantage of post content and rumor propagation structure as well as the dynamic changes of the heterogeneous graph when the rumor is spreading. Considering that the rumor content and propagation structure play essential roles in rumor classification, we combine the two factors into a union graph to obtain more rumor information details. Besides, GCN has the ability to tackle graph-structural data and learn higher-level representations, and the ordinary differential equation can acquire continuous-time dynamics on a complex graph. We design the dynamic module using GCN as the derivative network to get the dynamic representation of the heterogeneous graph. Finally, we conduct extensive experiments on two real-world datasets (Twitter15 and Twitter16), and the experimental results demonstrate that our proposed method outperforms mainstream models in rumor detection.

In the future, we will explore multimodal graphs, which can integrate more information such as user profiles, images, and semantic knowledge for rumor classification. Moreover, we will continue to dive into the study of social media’s rumor propagation dynamic system, building more efficient methods to explore the dynamic changes of rumor spreading to enhance the robustness of rumor classification.

Data Availability

The data underlying the results presented in the study are a public dataset from the paper “Detect Rumors in Microblog Posts Using Propagation Structure via Kernel Learning”

produced by Jing MA, which is available at https://www.dropbox.com/home/rumdetect2017/rumor_detection_acl2017.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Humanities and Social Sciences Project of the Ministry of Education (Grant no. 18YJA860020) and the Key Research and Development Program of Zhejiang Province, China (Grant no. 2019C01002).

References

- [1] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, “Fake news detection on social media,” *ACM SIGKDD explorations newsletter*, vol. 19, no. 1, pp. 22–36, 2017.
- [2] X. Zhou and R. Zafarani, “A survey of fake news,” *ACM Computing Surveys*, vol. 53, no. 5, pp. 1–40, 2020.
- [3] S. M. Alzanin and A. M. Azmi, “Detecting rumors in social media: a survey,” *Procedia Computer Science*, vol. 142, pp. 294–300, 2018.
- [4] P. Meel and D. K. Vishwakarma, “A temporal ensembling based semi-supervised ConvNet for the detection of fake news articles,” *Expert Systems with Applications*, vol. 177, Article ID 115002, 2021.
- [5] D. K. Vishwakarma and C. Jain, “Recent state-of-the-art of fake news detection: a review,” in *Proceedings of the 2020 International Conference for Emerging Technology (INCET)*, pp. 1–6, IEEE, Belgaum, India, June 2020.
- [6] X. Wang, Y. Li, C. Qiu, J. Li, and Y. Liu, “Research on reversal Model and Simulation of Network rumor Propagation under public health emergencies,” *Library and Information Service*, vol. 65, no. 19, pp. 4–15, 2021.
- [7] F. Yang, Y. Liu, X. Yu, and M. Yang, “Automatic detection of rumor on sina weibo,” in *Proceedings of the ACM SIGKDD workshop on mining data semantics*, pp. 1–7, Beijing China, August 2012.
- [8] S. Kwon, M. Cha, K. Jung, W. Chen, and Y. Wang, “Prominent features of rumor propagation in online social media,” in *Proceedings of the 2013 IEEE 13th international conference on data mining*, pp. 1103–1108, IEEE, Dallas, TX, USA, December 2013.
- [9] K. Popat, “Assessing the credibility of claims on the web,” in *Proceedings of the 26th International Conference on World Wide Web Companion*, pp. 735–739, Perth Australia, April 2017.
- [10] V. Qazvinian, E. Rosengren, D. Radev, and Q. Mei, “Rumor has it: identifying misinformation in microblogs,” in *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, pp. 1589–1599, Edinburgh UK, July 2011.
- [11] Z. Zhao, P. Resnick, and Q. Mei, “Enquiring minds: early detection of rumors in social media from enquiry posts,” in *Proceedings of the 24th international conference on world wide web*, pp. 1395–1405, Florence Italy, May 2015.
- [12] D. Varshney and D. K. Vishwakarma, “A review on rumour prediction and veracity assessment in online social network,”

- Expert Systems with Applications*, vol. 168, Article ID 114208, 2021.
- [13] J. Sampson, F. Morstatter, L. Wu, and H. Liu, "Leveraging the implicit structure within social media for emergent rumor detection," in *Proceedings of the 25th ACM international conference on information and knowledge management*, pp. 2377–2382, Indianapolis IN USA, October 2016.
 - [14] J. Ma, W. Gao, and K. F. Wong, "Detect rumors in microblog posts using propagation structure via kernel learning," in *Proceedings of the Association for Computational Linguistics*, Vancouver, Canada, July 2017.
 - [15] K. Wu, S. Yang, and K. Q. Zhu, "False rumors detection on sina weibo by propagation structures," in *Proceedings of the 2015 IEEE 31st international conference on data engineering*, pp. 651–662, IEEE, Seoul, Republic of Korea, April 2015.
 - [16] C. Castillo, M. Mendoza, and B. Poblete, "Information credibility on twitter," in *Proceedings of the 20th international conference on World wide web*, pp. 675–684, Hyderabad India, March 2011.
 - [17] J. Ma, W. Gao, P. Mitra et al., "Detecting rumors from microblogs with recurrent neural networks," in *Proceedings of the IJCAI*, pp. 3818–3824, New York NJ USA, July 2016.
 - [18] J. Ma, W. Gao, and K. F. Wong, "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, May 2018.
 - [19] J. Bruna, W. Zaremba, A. Szlam, and Y. LeCun, "Spectral networks and locally connected networks on graphs," 2013, <https://arxiv.org/abs/1312.6203>.
 - [20] F. Yu, Q. Liu, S. Wu, L. Wang, and T. Tan, "A convolutional approach for misinformation identification," in *Proceedings of the IJCAI*, pp. 3901–3907, Melbourne Australia, August 2017.
 - [21] F. Yu, Q. Liu, S. Wu, L. Wang, and T. Tan, "Attention-based convolutional approach for misinformation identification from massive and noisy microblog posts," *Computers & Security*, vol. 83, pp. 106–121, 2019.
 - [22] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," 2016, <https://arxiv.org/abs/1609.02907>.
 - [23] T. Bian, X. Xiao, T. Xu et al., "Rumor detection on social media with Bi-directional graph convolutional networks," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 1, pp. 549–556, New York, NY, USA, February 2020.
 - [24] Y. Chen, *Convolutional Neural Network for Sentence classification*, University of Waterloo, Waterloo, Canada, 2015.
 - [25] Y. Zhang and B. Wallace, "A sensitivity analysis of (and practitioners' guide to) convolutional neural networks for sentence classification," 2015, <https://arxiv.org/abs/1510.03820>.
 - [26] K. K. Shang, T. C. Li, M. Small, D. Burton, and Y. Wang, "Link prediction for tree-like networks," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 29, no. 6, Article ID 61103, 2019.
 - [27] Q. Huang, J. Yu, J. Wu, and B. Wang, "Heterogeneous graph attention networks for early detection of rumors on twitter," in *Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, IEEE, Glasgow, UK, July 2020.
 - [28] C. Zang and F. Wang, "Neural dynamics on complex networks," in *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, California, CA USA, July 2020.
 - [29] K.-K. Shang, W.-S. Yan, and X.-K. Xu, "Limitation of degree information for analyzing the interaction evolution in online social networks," *International Journal of Modern Physics C*, vol. 25, no. 10, Article ID 1450056, 2014.
 - [30] S. H. Rudy, S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Data-driven discovery of partial differential equations," *Science Advances*, vol. 3, no. 4, Article ID e1602614, 2017.
 - [31] T. Qin, K. Wu, and D. Xiu, "Data driven governing equations approximation using deep neural networks," *Journal of Computational Physics*, vol. 395, pp. 620–635, 2019.
 - [32] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Multistep neural networks for data-driven discovery of nonlinear dynamical systems," 2018, <https://arxiv.org/abs/1801.01236>.
 - [33] V. Shankar, G. D. Portwood, A. T. Mohan et al., "Learning non-linear spatio-temporal dynamics with convolutional Neural ODEs," in *Proceedings of the Third Workshop on Machine Learning and the Physical Sciences (NeurIPS 2020)*, December 2020.
 - [34] M. Raissi, A. Yazdani, and G. E. Karniadakis, "Hidden fluid mechanics: learning velocity and pressure fields from flow visualizations," *Science*, vol. 367, no. 6481, pp. 1026–1030, 2020.
 - [35] C. Yildiz, M. Heinonen, and H. Lähdesmäki, "ODE2VAE: deep generative second order ODEs with Bayesian neural networks," 2019, <https://arxiv.org/abs/1905.10994>.
 - [36] J. Choi, J. Jeon, and N. Park, "Learnable-Time ODE-based collaborative filtering," in *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, pp. 251–260, Queensland Australia, November 2021.
 - [37] E. Dupont, A. Doucet, and Y. W. Teh, "Augmented neural odes," 2019, <https://arxiv.org/abs/1904.01681>.
 - [38] R. Dayani, N. Chhabra, T. Kadian, and R. Kaushal, "Rumor detection in twitter: an analysis in retrospect," in *Proceedings of the 2015 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS)*, pp. 1–3, IEEE, Kolkata, India, December 2015.
 - [39] Y. Liu and Y. F. B. Wu, "Early detection of fake news on social media through propagation path classification with recurrent and convolutional networks," in *Proceedings of the Thirty-second AAAI conference on artificial intelligence*, Hilton, NO, USA, February 2018.
 - [40] H. Guo, J. Cao, Y. Zhang, J. Guo, and J. Li, "Rumor detection with hierarchical social attention network," in *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, pp. 943–951, Torino Italy, October 2018.
 - [41] D. Varshney and D. K. Vishwakarma, "Hoax news-inspector: a real-time prediction of fake news using content resemblance over web search results for authenticating the credibility of news articles," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 9, pp. 8961–8974, 2021.
 - [42] R. K. Kaliyar, A. Goswami, and P. Narang, "FakeBERT: fake news detection in social media with a BERT-based deep learning approach," *Multimedia Tools and Applications*, vol. 80, no. 8, pp. 11765–11788, 2021.
 - [43] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "Bert: pre-training of deep bidirectional transformers for language understanding," 2018, <https://arxiv.org/abs/1810.04805>.
 - [44] P. Meel and D. K. Vishwakarma, "HAN, image captioning, and forensics ensemble multimodal fake news detection," *Information Sciences*, vol. 567, pp. 23–41, 2021.
 - [45] D. Varshney and D. K. Vishwakarma, "A unified approach for detection of Clickbait videos on YouTube using cognitive

- evidences,” *Applied Intelligence*, vol. 51, no. 7, pp. 4214–4235, 2021.
- [46] B. Malhotra and D. K. Vishwakarma, “Classification of propagation path and tweets for rumor detection using graphical convolutional networks and transformer based encodings,” in *Proceedings of the 2020 IEEE Sixth International Conference on Multimedia Big Data (BigMM)*, pp. 183–190, IEEE, New Delhi, India, September 2020.
- [47] Y. Liu, M. Ott, N. Goyal et al., “Roberta: a robustly optimized bert pretraining approach. arXiv 2019,” 1907, <https://arxiv.org/abs/1907.11692>.
- [48] Y. Dou, K. Shu, C. Xia, P. S. Yu, and L. Sun, “User preference-aware fake news detection,” 2021, <https://arxiv.org/abs/2104.12259>.
- [49] R. T. Q. Chen, Y. Rubanova, J. Bettencourt, and D. Duvenaud, “Neural ordinary differential equations,” *Advances in Neural Information Processing Systems*, vol. 31, 2018.
- [50] T. Ma, H. Zhou, Y. Tian, and N. Al-Nabhan, “A novel rumor detection algorithm based on entity recognition, sentence reconfiguration, and ordinary differential equation network,” *Neurocomputing*, vol. 447, pp. 224–234, 2021.
- [51] D. Wang, X. Zhang, Y. Wan, D. Yu, G. Xu, and S. Deng, “Modeling sequential listening behaviors with attentive temporal point process for next and next new music recommendation,” *IEEE Transactions on Multimedia*, 2021.
- [52] J. Ramos, “Using tf-idf to determine word relevance in document queries,” in *Proceedings of the first instructional conference on machine learning*, vol. 242, no. 1, pp. 29–48, Washington D.C, USA, August 2003.
- [53] G. C. Layek, *An Introduction to Dynamical Systems and chaos [M]*, Springer, New Delhi, India, pp. 1–35, 2015.
- [54] S. Kwon, M. Cha, and K. Jung, “Rumor detection over varying time windows,” *PLoS One*, vol. 12, no. 1, Article ID e0168344, 2017.
- [55] J. Ma, W. Gao, Z. Wei, Y. Lu, and K.-F. Wong, “Detect rumors using time series of social context information on micro-blogging websites,” in *Proceedings of the 24th ACM international on conference on information and knowledge management*, pp. 1751–1754, Melbourne Australia, October 2015.
- [56] K. Tu, C. Chen, C. Hou, J. Yuan, J. Li, and X. Yuan, “Rumor2vec: a rumor detection framework with joint text and propagation structure representation learning,” *Information Sciences*, vol. 560, pp. 137–151, 2021.
- [57] J. R. Dormand, *Numerical Methods for Differential Equations: A Computational approach*, CRC Press, Boca Raton, FL, USA, 1996.