IRUGCN: A Graph Convolutional Network Rumor Detection Model Incorporating User Behavior

Shu Zhou†, Hao Wang,*†, Zhengda Zhou†, Haohan Yi†, Bin Shi†

南京大學
NANJING UNIVERSITY

Outline

1. Introduction
2. Methodology
3. Experiment and analysis
4. Conclusions

南京大學
NANJING UNIVERSITY

1. Introduction

- Research Background

With the popularity of social media, the spread of rumors on these platforms has become a serious problem.

- Research Question

Identifying these rumors by human are time-consuming and labor-intensive and also expensive.

- Research Gap

Current researches mainly focus on deep learning-based rumor detection methods, emphasizing rumor content and common user attributes, while ignoring the user behavior patterns.

- Our Solution

We propose the novel IRUGCN model, using graph convolutional networks to learn user representations, which not only considers the basic characteristics of users, but also incorporates the features of user behaviors.

2.1 Task definition

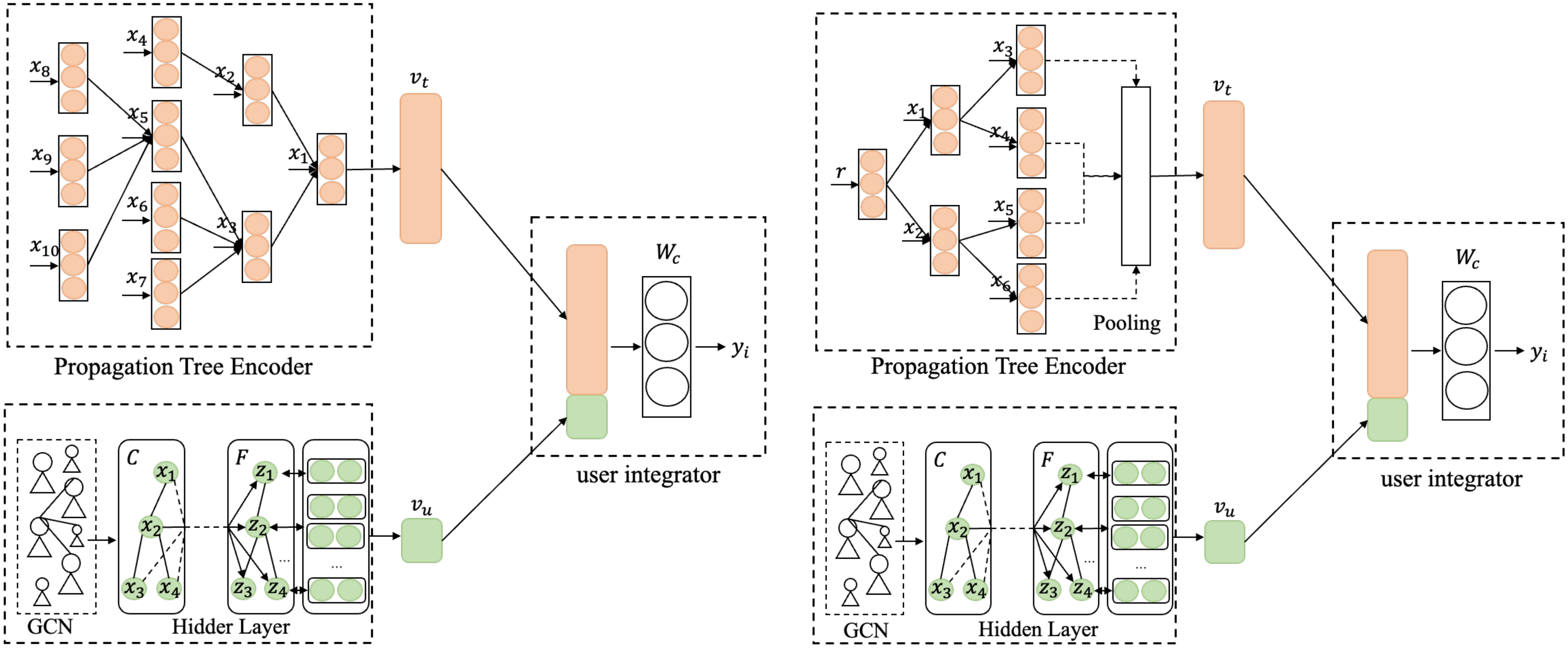
- Dataset

Rumor detection dataset is defined as several tuples $Eg = \{Eg_1, Eg_2, \dots, Eg_{|Eg|}\}$, where each tuple contains a set of declarations and a set of corresponding users i. $e. Eg = \{C_i, U_i\}$

- Objective

The objective of the rumor detection task is to construct a classifier for determining whether or not the declarations is a rumor. The classifier can be formalized as a function $f: C_i \rightarrow Y_i$,where Y_i is denoted as one of the following categories: non-rumor, false rumor, true rumor, and unconfirmed rumor.

2.2 Overall structure



(a) Low-Up Propagation Tree Encoders (b) Top-Down Propagation Tree Encoders

$$Loss\ fuction\text{: } L(Y,P) = -\frac{1}{m}\sum_{i=1}^{M}\sum_{k=1}^{K} y_{i,k}\log(p_{i,k}) + \lambda|\theta|$$

3.1 Experimental setup

Dataset：

- The study conducts experiments on two widely used Twitter datasets (Twitter15 and Twitter16), comprising 1381 and 1181 propagation trees respectively.
- Model performance is evaluated based on overall accuracy and F1 scores.

Baseline：

- BERT
- Transformer
- RvNN
- UMLARD
- DDGCN

3.2 Comparative analysis of methods

Table 2 Rumor Detection Performance Comparison Results

Model	Accuracy	Non-rumor	Fake rumor	Rumor	Unknown rumor
Twitter15 Dataset					
		F_1			
BERT	0.641	0.684	0.634	0.688	0.571
Transformer	0.708	0.695	0.728	0.759	0.653
RvNN	0.723	0.682	0.758	0.821	0.654
UMLARD	0.742	0.693	0.765	0.831	0.661
DDGCN	0.812	0.793	0.773	0.851	0.741
BU-IRUGCN	0.838	0.896	0.813	0.873	0.773
TD-IRUGCN	**0.852**	**0.799**	**0.873**	**0.931**	**0.809**
Twitter16 Dataset					
Model	Accuracy	Non-rumor	Fake rumor	Rumor	Unknown rumor
		F_1			
BERT	0.633	0.617	0.715	0.577	0.527
Transformer	0.718	0.723	0.712	0.799	0.659
RvNN	0.737	0.662	0.743	0.835	0.708
UMLARD	0.783	0.734	0.806	0.872	0.702
DDGCN	0.817	0.798	0.778	0.856	0.746
BU-IRUGCN	0.835	0.806	0.835	0.960	0.856
TD-IRUGCN	**0.873**	**0.816**	**0.856**	**0.970**	**0.856**

Note BU-IRUGCN denotes the use of a low-up propagation tree-structured encoder and TD-IRUGCN denotes the use of a top-down propagation tree-structured encoder

3.3 Encoder Effectiveness Analysis

- Methods that directly uses the user's statistical features (BU-Features and TD-Features): this method only relies on statistical data without further processing or extraction of deep features.
- Methods using fully connected layers (BU-SVD and TD-SVD): this method works by connecting user statistical features with a low-dimensional representation of the collocation matrix of user behavior.

Table 3 Performance of bottom-up propagation tree encoder combining different features on Twitter15 dataset

Twitter15					
Model	Accuracy	Non-rumor	Fake rumor	True Rumor	Unknown rumor
		F_1			
BU-Features	0.760	0.750	0.766	0.816	0.700
BU-SVD	0.775	0.763	0.773	0.859	0.707
BU-IRUGCN	**0.838**	**0.896**	**0.813**	**0.873**	**0.763**
Twitter16					
Model	Accuracy	Non-rumor	Fake rumor	True Rumor	Unknown rumor
		F_1			
BU-Features	0.763	0.665	0.805	0.866	0.687
BU-SVD	0.786	0.710	0.829	0.892	0.682
BU-IRUGCN	**0.835**	**0.806**	**0.835**	**0.960**	**0.736**

3.4 Ablation experiments

- Method 1: Remove the basic user features from the user encoder and use the top-down propagation tree structure encoder.
- Method 2: Remove the behavioral features in the user encoder and use the top-down propagation tree structure encoder.
- Method 3: Remove the basic user features in the user encoder and use the bottom-up propagation tree structure encoder.
- Method 4: Remove the behavioral features from the user encoder and use bottom-up propagation tree structure encoder.

Table 5 Ablation Study

Method	Twitter15	Twitter16
	$\Delta Average\ F_1$	
Method 1	-0.26	-0.23
Method 2	-0.47	-0.39
Method 3	-0.24	-0.19
Method 4	-0.43	0.36

3.5 Early Rumor Analysis

In order to validate the effectiveness of the proposed method in early rumor detection, we select three key time points (12h, 24h, and 36h) in rumor spreading process after the tweet is published and let model make decision based on the comment data before these time points.

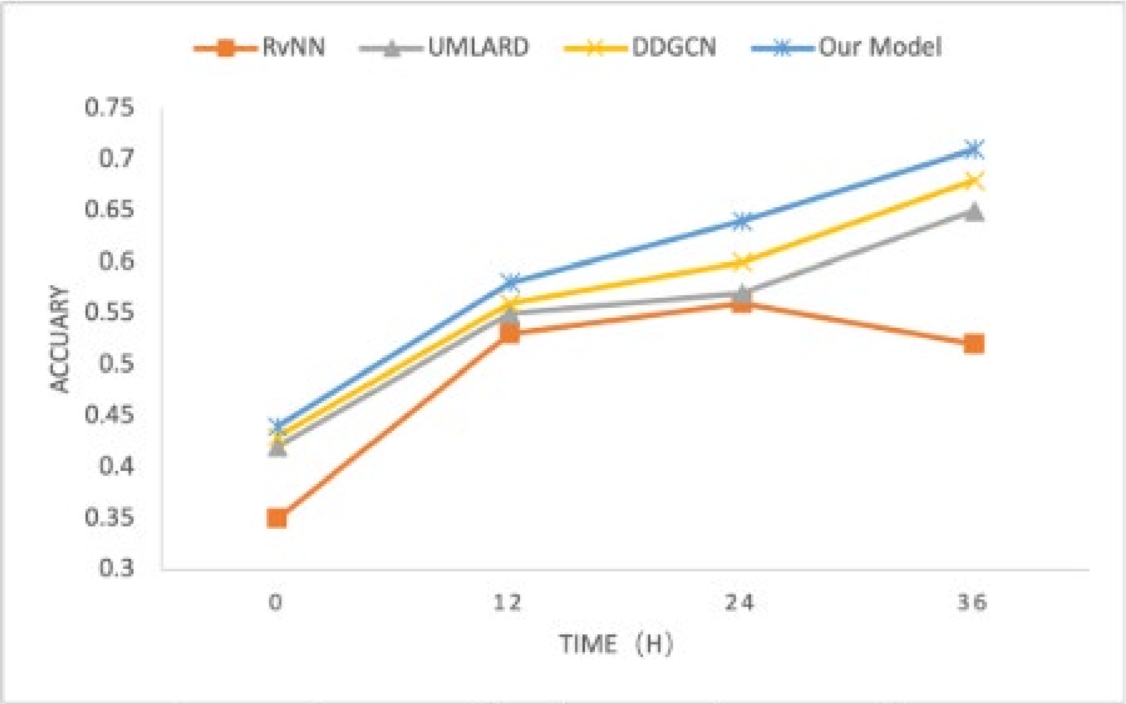


Figure 1: Early rumor model performance on the Twitter15 dataset

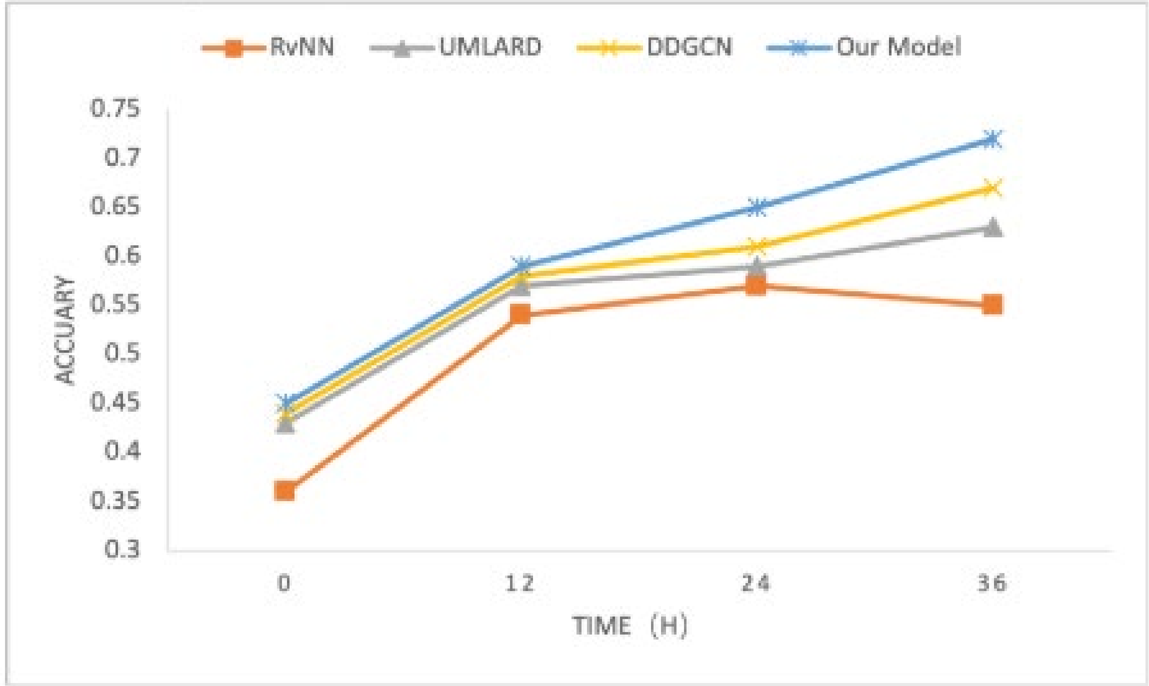


Figure 2: Early rumor model performance on the Twitter16 dataset

4. Summarize

- First, we introduced a novel graph neural network model for rumor detection which takes user behavior into account, making rumor detection more accurate and real-time.
- Second, the model integrates three components: a user encoder, a propagation tree structure encoder, and an integrator, facilitating a multi-dimensional analysis of rumors content, user behavior, and propagation process.
- Experiments results demonstrated our proposed model has obvious advantages and higher accuracy in early rumor detection compared to existing methods.
- The future research will focus on the user interactions. We plan to assign weights to the edges in the user graph so that the model could capture the more subtle relationships between users.

Thanks for listening

南京大学
NANJING UNIVERSITY