Fake News Detection with Context Awareness of the Publisher

Xingchen Ding, Chong Teng* , Donghong Ji
Key Laboratory of Aerospace Information Security and Trusted Computing, Ministry of Education,
School of Cyber Science and Engineering, Wuhan University, Wuhan, China
Email: {xingos,tengchong,dhji}@whu.edu.cn

***Abstract*—The spread of fake news is a significant social problem that can have disastrous impacts on various domains, such as politics and the economy. Therefore, detecting fake news has become a major concern. However, prior research has relied solely on news text to derive news representation, which is inadequate because different news items under the same publisher are interconnected. To address this limitation, we propose an innovative approach called the Publisher-oriented Multi-view Graph Model (PMGM) that leverages the context awareness of the publisher to detect fake news. Our approach enriches the news representation by incorporating publisher profiles and text style features extracted from the news. Specifically, we construct a multi-view graph that encodes various relationships between news items from the same publisher, such as news topics and occasions in which they were released. Furthermore, we leverage a multi-layer Graph Convolutional Network in conjunction with jumping knowledge networks to model the multi-view graph and produce a publisher-oriented contextualized representation of news. Experimental results on two widely used fake news datasets, namely LIAR and Weibo21, demonstrate the effectiveness of our approach. Specifically, the PMGM model outperforms the state-of-the-art methods significantly. Overall, our proposed model unifies various heterogeneous features and information related to news based on a publisher-oriented approach, thereby offering a novel idea to enhance fake news detection.**

***Index Terms*—Fake News Detection, Multi-View Graph, Representation Learning, Relation Graph,**

I. Introduction

Currently, the proliferation of fake news poses a significant threat to the reliability and veracity of news. The public's trust in the British government during the "Brexit" referendum and the fairness of the 2016 U.S. presidential election have been greatly undermined by the spread of fake news [1], [2]. As a result, there has been a growing interest in the NLP community in the development of fake news detection systems that can automatically assess the authenticity of a given news text [3], [4], [5], [6], [7].

Early research in this area focused on the manual engineering of features [3], [4]. Initially, researchers created comprehensive sets of hand-crafted features based on news content, user profiles, and news propagation paths. They then trained machine-learning classifiers to distinguish between true and false news. However, recent studies [6], [7], [8] have leveraged the success of deep learning and applied various neural network models, such as Convolutional Neural Networks (CNN)[9], Graph Convolutional Networks (GCN)[10], and BERT [11], to learn distinctive features from news text and identify fake news.

*Corresponding author
DOI reference number: 10.18293/SEKE2023-061

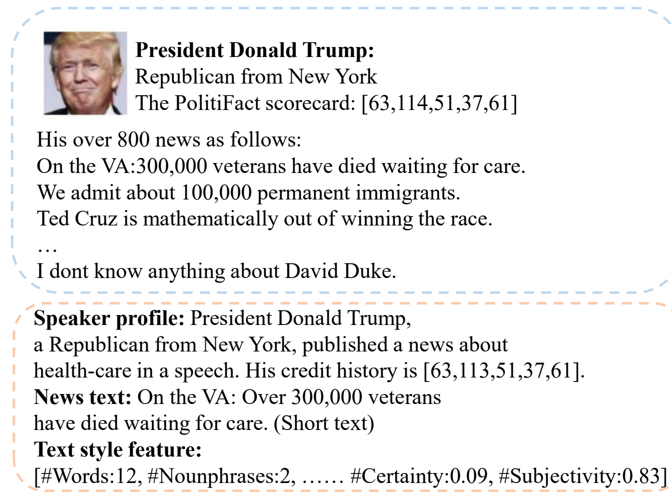


Fig. 1. Donald Trump's profile and published news (blue rectangle) as well as the features used in this paper (red rectangle).

Despite these advances, previous works have limited the representation of news to the news text, leading to insufficient utilization of other news-related information, such as publisher profiles and text style features. It is intuitive that the authenticity of news is related to these factors. For instance, as shown in Figure 1, Donald Trump's profile includes the PolitiFact scorecard of his news published before, with 70% of his news being false and only 30% being true. Furthermore, news with a subjectivity in text style features above 0.7 is likely to be false [12]. Another limitation of prior research is that it only employs a single-view graph based on one of multiple relationships between news for fake news detection [13], [14], [15]. However, it is intuitive that the topics of news and the occasions in which the news is published can also serve as heuristic factors for detecting fake news. For instance, during interviews, Trump only published 60% true news, while he seldom told the truth when publishing news related to the election. Therefore, we assume that news topic and social occasion information can have a significant impact on the detection of fake news.

To address the aforementioned limitations, we propose a novel multi-view graph model that is oriented towards the publisher (**PMGM**). Our model extracts 40 latent textual features as text style features given a piece of news text. It encodes the words in the news text and publisher profile using BERT, and then combines these features to represent the news text. Furthermore, we propose a multi-view news graph that is oriented towards the publisher. The nodes of the graph are news representations, and the edges model the relationships between news in different views, such as topics and occasions. The news representations are updated using Graph Convolutional Networks (GCN) and Jumping Knowledge Networks (JK-Nets) [16]. The representations from different views are fused using the attention mechanism [17]. Finally, a neural classifier is used to predict the labels of all news jointly. To demonstrate the effectiveness of our model, we conduct experiments on two fake news datasets, LIAR and Weibo21. The results show that PMGM significantly outperforms state-of-the-art methods.

II. Related Work

A. Feature-Based Methods

Feature-based fake news detection methods rely on either hand-crafted linguistic or embedding features to extract information about the writing style or language used in news text [5], [18], [19], [8]. For instance, Castillo et al. [3] introduced a decision tree-based model that utilized an extensive range of features for identifying fake news on Twitter. Similarly, Yang et al. [4] extracted a broad set of features from microblog data to train a classifier for automatically detecting fake news on Sina Weibo. Rubin et al. [20] provided a conceptual overview of satire and humor and illustrated the unique features of satirical news to detect potentially misleading information. Wu et al. [21] used a propagation structure composed of 23 features in the hybrid support vector machines (SVM) method. Alternatively, Wu et al. [22] proposed a machine learning model that relied on time series fitting of tweet volume time characteristics. However, these methods heavily rely on manual feature engineering and may not capture enough features to achieve satisfactory performance when dealing with short news articles.

B. Relation-Based Methods

As the fields of deep learning and graph neural networks continue to advance, relation-based methods that exploit certain connections between news articles have come to dominate this domain [13], [14], [15]. Long et al. [5] utilized publisher profiles as a means of representing the attention factors between news articles to propose a hybrid LSTM model for detecting fake news. Karimi et al. [18] combined information from multiple sources to discriminate between different degrees of fake news by taking into account the relationships between them. Hu et al. [13] proposed a graph that incorporates publisher profiles for fake news detection using multi-depth graph convolutional networks (M-GCN). Mendoza et al. [23] analyzed the topology of retweeting networks and identified differences between rumor diffusion patterns on Twitter and traditional news platforms. Li et al. [24] combined objective information with subjective factors for rumor detection. Kwon et al. [25] introduced the Periodic External Shocks (PES) model, which combines a set of linguistic features with network structure to identify rumors. However, all of these methods utilize only homogenous relationships between news articles from a single perspective, without considering the possibility of heterogeneous relationships.

III. Method

A. Model Overview

We present a novel approach for detecting fake news, which we call the Publisher-Oriented Multi-View Graph Model (**PMGM**). As shown in Figure 2, our framework comprises three key components. Firstly, our model extracts 40 latent textual features as text style features from a given news article and its corresponding publisher profile. We encode both the news article and publisher profile using BERT, as detailed in Section III-B, to obtain the enhanced contextualized representations of each news. These features are then combined to represent the news. Secondly, we construct a publisher-oriented multi-view graph, where nodes represent the news representations and edges are created based on the associations of topics and occasions between news articles from the same publisher, as discussed in Section III-C. Finally, we apply an attention mechanism to fuse the features of different views in the graph, and use a neural classifier to determine the labels of all news articles.

B. Enhance News Representation

We amalgamate the three characteristics of news text, publisher profile, and text style feature to form the publisher-oriented representation of a news item. Subsequently, we delineate the processing and embedding of the aforementioned features separately.

Encoding of News Text. To encode the news text, we utilize BERT as the embedding layer to extract the output vectors of all word tokens. The resulting global vector t_i corresponding to the [CLS] token represents the i_{th} news.

Encoding of Publisher Profile. We use various publisher profile information, including *party*, *publisher name*, *home state*, *credit history*, *social occasion* and *topic*, to enhance the performance of the fake news classifier. For discrete publisher profile information, we directly use their encoded representation. For instance, the credit history [23,12,22,43,61] is encoded as $s_i^1 = [23, 12, 22, 43, 61]$. For text-format publisher profile information, we reorganize them using the template *job publisher name*, a *party* from *home state*, published a news about *topic* in a *social occasion* and then utilize BERT as the embedding layer to transform the information into feature vectors s_i^2. Finally, we concatenate s_i^1 and s_i^2 to obtain s_i, which represents the publisher profile of the i_{th} news.

Extraction of Text Style Features. According to the definition of text style [26], a set of quantifiable characteristics (such as machine learning features) can effectively represent text content. Given a news text N to be verified, we represent it as a set $d_i = \{d_i^1, d_i^2, d_i^3....d_i^k\}$ of k text style features, where d_i^k typically takes the form of a number. To adequately capture

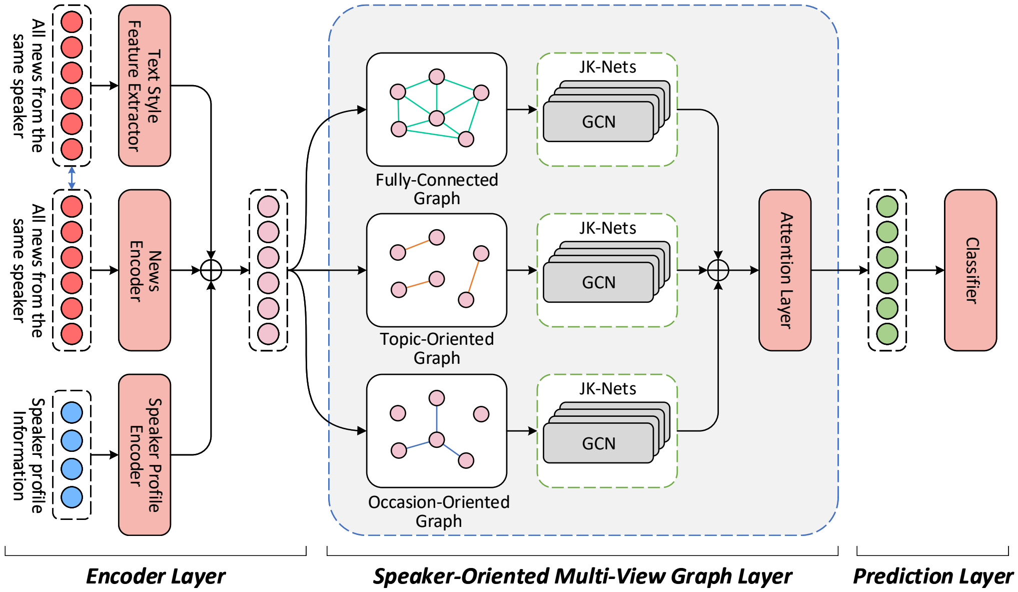


Fig. 2. The proposed PMGM framework for fake news detection. Pink nodes mean enhanced news representations and green nodes mean publisher-oriented context-aware news representations.

TABLE I
TEXT STYLE FEATURE TEMPLATES. '#' DENOTES THE AMOUNT.

Category	Feature
Quantity	#Characters; #Words; #Nounphrases; #Paragraphs; #Sentences
Complexity	Average #characters per word; Average #words per sentence; Average #clauses per sentence; Average #punctuations per sentence
Uncertainty	#Modal verbs; #Certainty terms; #Generalizing terms; #Tentative terms; #Numbers and quantifiers; #Question marks
Subjectivity	#Biased lexicons; #Subjective verbs; #Report verbs; #Factive verbs
Non-immediacy	#Self reference: 1st person singular pronouns; #Group reference: 1st person plural pronouns; #Other reference: 2nd and 3rd person pronouns; #Quotations
Sentiment	#Positive words; #Negative words; #Anxiety/angry/sadness words; #Exclamation marks; #Content sentiment polarity
Diversity	#unique words or terms; #unique content words; #unique function words; #Unique nouns/verbs/adjectives/adverbs
Specificity	#Temporal/spatial ratio; #Sensory ratio; #Causation terms; #Exclusive terms
Readablity	#Flesch-Kincaid and Gunning-Fog index

and represent the style of news text, we synthesized various fake news detection papers to obtain 40 features [26], [27] as shown in Table I. We then include these features in our news representation. Finally, the representation of each news h_i is summarized as $h_i = [t_i||s_i||d_i]$, where $|$ denotes concatenation.

C. Publisher-oriented Multi-view Graph Modeling

Multi-view Graph Design. The news published by a single entity covers diverse topics and occasions, posing a challenge in explicitly modeling the correlation between all news solely based on a single viewpoint [28]. To address this, we propose a publisher-oriented multi-view news graph that leverages distinct publisher perspectives. We establish edges among news nodes of the same publisher based on shared topic or social occasion. We create three graphs, namely the publisher full-connected graph, topic-oriented graph, and occasion-oriented graph. The occasion devotes social occasions, such as TV shows, interviews, and election campaigns, where news is published. For the publisher full-connected graph, we define the sparse adjacency matrix A^1 as follows:

$$A_{ij}^1 = \begin{cases} 1 \text{ if } i,j \text{ from same publisher} \\ 0 \text{ otherwise} \end{cases} \tag{1}$$

For the topic-oriented graph, we define the sparse adjacency matrix A^2 as follows:

$$A_{ij}^2 = \begin{cases} 0.5 \text{ if } i,j \text{ have the same topic from same publisher} \\ 0 \text{ otherwise} \end{cases} \tag{2}$$

For the occasion-oriented graph, we define the sparse adjacency matrix A^3 as follows:

$$A_{ij}^3 = \begin{cases} 0.2 \text{ if } i,j \text{ have the same occasion from same publisher} \\ 0 \text{ otherwise} \end{cases} \tag{3}$$

where i and j are the indices of different news.

Graph Embedding. Graph Convolutional Networks (GCNs) have shown significant promise in achieving generalization in various tasks. Our work builds upon the GCN module. At the ith layer, the GCN module takes the graph adjacency matrix A_i and hidden representation matrix $H_i = \{h_1, h_2, ..., h_{|\mathcal{D}|}\}$ with $|\mathcal{D}|$ news articles as input. The GCN module then outputs a hidden representation matrix $H_{i+1} \in \mathbb{R}^{n_i * d_{i+1}}$, which can be described as:

$$H_{i+1} = \text{ReLU}\left(\tilde{D}_i^{-\frac{1}{2}}\tilde{A}_i\tilde{D}_i^{-\frac{1}{2}}H_i\theta_i\right) \tag{4}$$

where adjacency matrix with self-loop $\tilde{A}_i = A_i + I$, $\tilde{D}_i$ is the degree matrix of $\tilde{A}_i$, and $\theta_i \in \mathbb{R}^{d_i * d_{i+1}}$ is a trainable weight matrix.

To enhance the structure-aware representation of a node, we utilize a combination of GCN and jumping knowledge networks (JK-Nets) based on the approach proposed by Xu et al. [16]. JK-Nets involve the direct transfer of each layer's representation to the final layer of the network, rather than solely passing it to the next layer of the convolutional network. This facilitates the aggregation of information from distinct receptive fields at the last layer, and allows the training process to maximize features by determining the receptive field size for each node. We obtain the output h_i^l for each graph node as follows:

$$h_i^l = \max(h_1, ..., h_i) \tag{5}$$

where h_i is the node representation of the ith GCN layer for the node h of H_i.

Learn Publisher-oriented Representation. After encoding both the node features and the graph structure of a multi-view graph in an end-to-end manner, we aggregate the multi-view information using an attention mechanism to form an updated representation. Specifically, we assign an attention score u_i (where $1 \leq i \leq m$) to the node representation of each view, which is then normalized using the softmax function. The publisher-oriented representation h_i^u for each news item is obtained by computing the weighted summation of the individual view representations h_i^l, where the weights are determined by the attention scores.

$$u_i = \tanh\left(W_i h_i^l + b_i\right); \alpha_i = \frac{\exp\left(u_i\right)}{\sum_{l=1}^{m} \exp\left(u_l\right)}; h_i^u = \sum_{i=1}^{m} \alpha_i h_i^l \tag{6}$$

D. Fake News Prediction

Research by Dou et al. [29] has shown that incorporating updated news representation with the original news can enhance the performance of fake news detection. Prior to inputting the final representation into the classifier, we concatenate the [CLS] representation of the original news with the publisher-oriented news representation. Subsequently, we apply a softmax classifier to predict the truthfulness label of the news.

$$\hat{y}_i = \mathrm{softmax}\left(\mathrm{ReLU}\left(\left[h_i^u \| o_i\right] W_{tp} + b_{tp}\right)\right) \tag{7}$$

Where W_{tp} and b_{tp} represent the parameters of the output layer, h_i^u and o_i represent the updated news representation and the [CLS] representation of the original news, respectively. Our model is trained using the cross-entropy loss during the training phase, which can be formalized as:

$$\mathcal{L} = -\sum_{i \in \mathcal{D}} y_i \ln \hat{y}_i \tag{8}$$

Where y_i represents the ground-truth label for the truthfulness of the i_{th} news article, and $\hat{y}_i$ represents the predicted distribution for the truthfulness label of the same article. Our objective is to minimize the loss function $\mathcal{L}$ for the purpose of detecting fake news.

IV. Experimental Settings

A. Dataset

We evaluated our model using two datasets, LIAR [6] and Weibo21 [30], which contain instances of fake news. The LIAR dataset includes 12,800 human-labeled short news items with six fine-grained labels that indicate the degree of truthfulness, including *pants-fire*, *false*, *barely-true*, *half-true*, *mostly-true*, and *true*. The dataset's label distribution is as follows: 1,050 for *pants-fire*, and a range of 2,063 to 2,638 for the other labels. On the other hand, the Weibo21 dataset is a Chinese multi-domain dataset with 4,488 fake news items and 4,640 real news items. These datasets are unique in that they include several metadata features, such as the topic, publisher, job, state, party, and total credit history count of the news publisher.

TABLE II
Top 5 publishers, news topics, occasions.

Top-5 Publishers	Num.	Top-5 Topics	Num.	Top-5 Occasions	Num.
Barack Obama	611	healthcare	474	news release	309
Donald Trump	343	taxes	356	interview	286
Hillary Clinton	297	education	309	press release	282
Mitt Romney	212	elections	304	speech	259
John Mccain	189	immigration	303	TV ad	222

As standard practice in machine learning, we split the datasets into training (80%), validation (10%), and testing (10%) sets. The LIAR dataset has 3,308 publishers, 144 news topics, and 302 occasions. To provide an overview of the publishers, topics, and occasions in this dataset, we present the top-5 most frequent publishers, topics, and occasions in Table II.

B. Implementation Details

During the text processing stage, the initial step involves cleansing the text information by eliminating redundant expressions and symbols, standardizing the case, and so forth. In this study, we utilized Bert-base to acquire 768-dimensional embeddings for each news. The hidden unit dimensions in GCN were established as [768, 768], while the learning rate was 0.001 and the number of GCN layers was 6. A dropout rate of 0.5 was specified. To optimize all parameters, we employed the Adam optimizer, coupled with a weight decay strategy, to train the model for 80 epochs. To ensure a level playing field for comparisons, we adopted the same evaluation metrics that were utilized in previous research, namely Accuracy and F1-measure (F1).

C. Baselines

We compared our proposed models with existing fake news detection models, including state-of-the-art models, on both the LIAR dataset and Weibo21 dataset. The models that we compared with are as follows: 1)**CNN-WangP [6]:** A hybrid CNN that integrates both text and contextual information to detect fake news. 2)**MMFD [18]:** A multi-source, multi-class fake news detection model that employs multiple sources of information to detect fake news across various classes. 3)**LSTM-Attention [5]:** A hybrid LSTM that accounts for word importance using an attention mechanism. 4)**FT+BERT [19]:** A fine-tuning technique based on the BERT pre-trained language model, which we utilized in our study. 5)**FakeBERT [8]:** A model that combines various parallel blocks of a single-layer deep CNN, each with different kernel sizes and filters, with the BERT. 6)**M-GCN [13]:** A semi-supervised fake

TABLE III

ACCURACIES AND F1S ON TWO TEST SETS. ABBREVIATIONS: SOURCE NEWS TEXT (**ST**), TEXT STYLE FEATURES (**TSF**), SPEAKER PROFILE FEATURES (**SF**), RELATIONSHIP BETWEEN NEWS (**RN**), ATTENTION MECHANISM (**AT**) AND MULTI-VIEW GRAPH (**MV**). ✓INDICATES THAT THE METHOD INCLUDES THAT FEATURE.

Dataset							Liar (6 classes)		Weibo21 (2 classes)	
Method	**ST**	**TSF**	**SF**	**RN**	**AT**	**MV**	**Accurary**	**Macro-F1**	**Accurary**	**Macro-F1**
CNN-WangP	✓		✓				0.274	0.265	0.774	0.818
MMFD	✓		✓				0.388	0.376	0.803	0.827
LSTM-Attention	✓		✓		✓		0.393	0.401	0.814	0.835
FT+BERT	✓		✓				0.423	0.451	0.839	0.856
FakeBERT	✓		✓				0.445	0.473	0.843	0.868
M-GCN	✓		✓	✓	✓		0.471	0.478	0.854	0.879
MTFake	✓		✓	✓			0.476	0.495	0.869	0.893
MMFake	✓		✓	✓		✓	0.492	0.531	0.872	0.891
PMGM w/o BERT	✓	✓	✓	✓	✓	✓	0.532	0.569	0.883	0.913
PMGM	✓	✓	✓	✓	✓	✓	0.543	0.582	0.894	0.921

news detection method that leverages text content as node features and publisher profiles to build a graph. 7)**MTFake [14]:** A multitask learning model that categorizes news articles collected from the web as either fake or not. 8)**MMFake [15]:** A multitopic and multitask fake news detection model that addresses the limitations of fusing different topics. 9)**PMGM:** Our proposed publisher-oriented multi-view graph modeling (PMGM) framework.

V. RESULTS AND ANALYSIS

A. Comparisons with Baselines

Table III presents the primary experimental results, along with the inter-model differences. Notably, the proposed PMGM model demonstrates significantly better performance than the state-of-the-art model on both the LIAR dataset and Weibo21 dataset. It is clear from the table that the relation-based methods (M-GCN, MTFake, MMFake, and PMGM) significantly outperform the feature-based methods (CNN-WangP, MMFD, LSTM-Attention, FT+BERT, and FakeBERT). This is largely attributed to the relation-based methods' ability to learn semantically rich representations of news, which enable them to capture more effective features. This underscores the importance of studying relation-based methods for fake news detection.

Specifically, our proposed PMGM method outperforms the MTFake and MMFake methods for two main reasons. Firstly, these methods do not delve into the representations of news texts, which are crucial in short news texts. Secondly, while MMFake only considers a relationship between news, PMGM considers both publisher-oriented multiple relations between news. This highlights the effectiveness of enhancing the representation of short news texts and employing a publisher-oriented multi-view graph model for fake news detection.

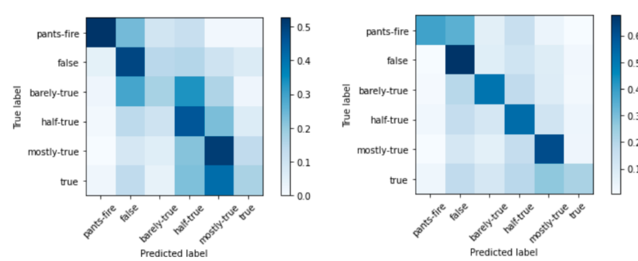
B. Ablation Studies

We conducted a series of ablation studies on key parts of the MVAN in order to determine the relative importance of each module, using the LIAR dataset. The experimental results of this comparison are presented in Table IV. Our findings indicate that when the PMGM model removes the publisher profiles, performance drops by approximately 25%. This can be attributed to the fact that publisher profiles contain credit history, which is a statistical dataset collected from previous statements of publishers and not readily available. Furthermore, when compared to MMFake without credit history, our model shows an improvement in accuracy of about 1.67%. We also discovered that removing each of the publisher-oriented multi-view components led to a drop in performance of approximately 2.4%, 1.7% and 0.6% respectively, demonstrating that multi-view mechanisms have a significant impact on model performance. Additionally, we observed that not using the [CLS] of original news led to a 1% drop in model performance. This indicates that combining the updated news representation with the original news can significantly enhance the performance of fake news detection.

TABLE IV

ABLATION ANALYSIS FOR PMGM.

Method	Accurary	Macro-F1
PMGM	**0.543**	**0.582**
w/o Publisher profiles	0.291	0.293
w/o Text style features	0.528	0.514
w/o Publisher view	0.498	0.518
w/o Topic view	0.526	0.531
w/o Occasion view	0.537	0.548
w/o JK-Nets	0.531	0.569



(a) Single-view graph modeling (b) Multi-view graph modeling

Fig. 3. Confusion matrix comparisons using single-view and multi-view in our PMGM.

C. Effect Of the Multi-View Graph Modeling

To explore the impact of a multi-view module on fake news detection, we demonstrate the significance of the module in Figure 3. The results indicate that the model achieves higher accuracy in classifying false, barely-true, half-true, and mostly-true news when the multi-view module is utilized, and the model's performance in correctly classifying labels declines when the module is removed. Hence, our approach, PMGM, leverages a publisher-oriented multi-view graph to aid the detection task, as the limited information present in short news content makes it difficult to obtain sufficient representations during model learning, leading to poor model performance.

VI. Conclusion

In this paper, we proposes a novel deep learning model, PMGM, for detecting fake news. The model combines a rich news representation with a multi-view of the publisher to capture important hidden clues and information in both the news text and its publisher. Our evaluation using two public datasets demonstrates that PMGM outperforms existing methods in fake news detection. Furthermore, we anticipate that utilizing the publisher-oriented multi-view graph model will also prove advantageous in other text classification tasks, such as sentiment and topic classification. In future work, we intend to expand our research in the following areas: (1) constructing a larger and more current dataset that includes additional publisher and user profiles and propagation data, (2) incorporating publisher and user information using more sophisticated methods, and (3) integrating additional publisher information into PMGM to achieve more robust evaluations of the fake news model.

Acknowledgements

This work is supported by the National Key Research and Development Program of China No. 2017YFC1200500, the Research Foundation of Ministry of Education of China No. 18JZD015.

References

[1] D. Pogue, "How to stamp out fake news," *Scientific American*, vol. 316, no. 2, pp. 24–24, 2017.

[2] R. Zafarani, X. Zhou, K. Shu, and H. Liu, "Fake News Research: Theories, Detection Strategies, and Open Problems," in *Proceedings of SIGKDD 2019*, pp. 3207–3208, ACM, 2019.

[3] 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, 2011.

[4] 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, 2012.

[5] Y. Long, Q. Lu, R. Xiang, M. Li, and C. Huang, "Fake news detection through multi-perspective speaker profiles," in *Proceedings of the Eighth International Joint Conference on Natural Language Processing, IJCNLP 2017, Taipei, Taiwan, November 27 - December 1, 2017, Volume 2: Short Papers*, pp. 252–256, 2017.

[6] W. Y. Wang, "" liar, liar pants on fire": A new benchmark dataset for fake news detection," *arXiv preprint arXiv:1705.00648*, 2017.

[7] M. H. Goldani, S. Momtazi, and R. Safabakhsh, "Detecting fake news with capsule neural networks," *Applied Soft Computing*, vol. 101, p. 106991, 2021.

[8] 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.

[9] T. He, W. Huang, Y. Qiao, and J. Yao, "Text-attentional convolutional neural network for scene text detection," *IEEE TIP*, vol. 25, no. 6, pp. 2529–2541, 2016.

[10] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," *arXiv preprint arXiv:1609.02907*, 2016.

[11] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.

[12] C. L. M. Jeronimo, L. B. Marinho, C. E. Campelo, A. Veloso, and A. S. da Costa Melo, "Fake news classification based on subjective language," in *Proceedings of iiWAS 2019*, pp. 15–24, 2019.

[13] G. Hu, Y. Ding, S. Qi, X. Wang, and Q. Liao, "Multi-depth graph convolutional networks for fake news detection," in *CCF NLPCC*, pp. 698–710, Springer, 2019.

[14] S. Sridhar and S. Sanagavarapu, "Fake news detection and analysis using multitask learning with bilstm capsnet model," in *2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, pp. 905–911, IEEE, 2021.

[15] C. Liang, Y. Zhang, X. Li, J. Zhang, and Y. Yu, "Fudfend: Fuzzy-domain for multi-domain fake news detection," in *Natural Language Processing and Chinese Computing: 11th CCF International Conference, NLPCC 2022, Guilin, China, September 24–25, 2022, Proceedings, Part II*, pp. 45–57, Springer, 2022.

[16] K. Xu, C. Li, Y. Tian, T. Sonobe, K.-i. Kawarabayashi, and S. Jegelka, "Representation learning on graphs with jumping knowledge networks," in *Proceedings of ICML 2018*, pp. 5453–5462, PMLR, 2018.

[17] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[18] H. Karimi, P. Roy, S. Saba-Sadiya, and J. Tang, "Multi-source multi-class fake news detection," in *Proceedings of COLING 2018*, pp. 1546–1557, 2018.

[19] A. Malakhov, A. Patruno, and S. Bocconi, "Fake news classification with bert.," in *MediaEval*, 2020.

[20] V. L. Rubin, N. Conroy, Y. Chen, and S. Cornwell, "Fake news or truth? using satirical cues to detect potentially misleading news," in *Proceedings of CADD 2016*, pp. 7–17, 2016.

[21] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and S. Y. Philip, "A comprehensive survey on graph neural networks," *IEEE TNNLS*, vol. 32, no. 1, pp. 4–24, 2020.

[22] K. Wu, S. Yang, and K. Q. Zhu, "False rumors detection on sina weibo by propagation structures," in *Proceedings of ICDM 2015*, pp. 651–662, IEEE, 2015.

[23] M. Mendoza, B. Poblete, and C. Castillo, "Twitter under crisis: Can we trust what we rt?," in *Proceedings of the first workshop on social media analytics*, pp. 71–79, 2010.

[24] J. Li, S. Ni, and H.-Y. Kao, "Meet the truth: Leverage objective facts and subjective views for interpretable rumor detection," *arXiv preprint arXiv:2107.10747*, 2021.

[25] S. Kwon, M. Cha, K. Jung, W. Chen, and Y. Wang, "Prominent features of rumor propagation in online social media," in *Proceedings of the ICDM*, pp. 1103–1108, IEEE, 2013.

[26] X. Zhou, R. Zafarani, K. Shu, and H. Liu, "Fake news: Fundamental theories, detection strategies and challenges," in *Proceedings of WSDM 2019*, pp. 836–837, 2019.

[27] M. Siering, J.-A. Koch, and A. V. Deokar, "Detecting fraudulent behavior on crowdfunding platforms: The role of linguistic and content-based cues in static and dynamic contexts," *JMIS*, vol. 33, no. 2, pp. 421–455, 2016.

[28] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake news detection on social media: A data mining perspective," *ACM SIGKDD explorations newsletter*, vol. 19, no. 1, pp. 22–36, 2017.

[29] Y. Dou, K. Shu, C. Xia, P. S. Yu, and L. Sun, "User preference-aware fake news detection," in *Proceedings of SIGIR 2021*, pp. 2051–2055, 2021.

[30] Y. Zhu, Q. Sheng, J. Cao, Q. Nan, K. Shu, M. Wu, J. Wang, and F. Zhuang, "Memory-guided multi-view multi-domain fake news detection," *IEEE Transactions on Knowledge and Data Engineering*, 2022.