MTF-GCN: A MULTIVIEW TEMPORAL FUSION GRAPH CONVOLUTIONAL NETWORK FOR RUMOR DETECTION AND FACT-CHECKING ON SOCIAL MEDIA

Tieu Phung Mai Suong *, Nguyen Thi Thuy A

Faculty of Information Technology,
Ho Chi Minh City University of Foreign Languages – Information Technology, Vietnam

suongtpm@huflit.edu.vn, antt@huflit.edu.vn

ABSTRACT – Rumors and low-reliability sources spread very quickly on social media and can have serious impacts on society. Many existing methods rely heavily on textual content or static propagation structures, thus limiting their ability to capture the dynamic interaction between information and users over time. In this study, we propose MTF-GCN, a multi-perspective time-based consolidation model for rumor detection and information verification at the document level. Each event is represented by two dynamic graphs: (i) propagation graphs constructed from reply and reshare relationships, and (ii) user graphs based on profile and interaction characteristics. The diffusion process is divided into three time phases of early, intermediate, and late, which are then processed using graph convolution layers and synthesized through a compact time-consolidation mechanism. The representations obtained from the propagation branch and the user branch are further combined to perform binary classification. Experiments on Weibo, Pheme, and ReCOVery datasets showed that MTF-GCN achieved high accuracy (97.75%, 95.22%, and 93.10%, respectively) and was particularly strong in recalling untrusted content. These results suggest that multi-perspective spatio-temporal modeling, combined with a simple time-based consolidation mechanism, is effective and suitable for early-warning and misinformation-monitoring systems.

Keywords – Rumor detection, fact-checking, social media, graph neural networks, multiview spatio-temporal modeling, user interaction

I. INTRODUCTION

Social media has become the main channel for receiving and sharing information, but at the same time, it has also increased the speed of spreading rumors and unreliable sources. Misinformation can distort public opinion, erode public trust, and lead to serious consequences in sensitive contexts such as crises or elections. Therefore, automatically detecting rumors and assessing the credibility of news remains an important but challenging task, because text on social networks is often short, noisy, and very diverse in expression, making it difficult for models, based only on content, to fully capture the context. Recent studies show that authenticity is not only determined by content, but also reflected in how information spreads and how users interact with it. The structure and pacing of responses, reshares, and comments indicate whether an assertion is being passively amplified or actively debated. Graph-based models, which represent events through propagation trees and user engagement networks, have proven highly effective in detecting rumors, especially when integrating spatio-temporal perspectives. However, most existing methods are focused on microblog data and rely on static graphs or complex consolidation mechanisms, which limit their applicability for assessing the reliability of news at the document level. This motivates the need for a simpler and more unified framework that can model the temporal evolution of both information propagation and user behavior.

In this work, we propose MTF-GCN (Multiview Temporal Fusion Graph Convolutional Network), which represents each event using two dynamic graphs, one for message propagation and one for user interactions, segments the diffusion process into temporal stages, and integrates them through a lightweight fusion mechanism in a shared latent space. Experiments on Weibo, Pheme, and ReCOVery show that MTF-GCN achieves competitive or state-of-the-art performance, with particularly high recall on unreliable content, making it well-suited for early warning in misinformation monitoring systems.

II. RELATED WORK

A. CONTENT-BASED RUMOR DETECTION

Early work on rumor detection framed the task as text classification at the post- or event-level, relying on hand-crafted linguistic features such as n-grams, sentiment cues, punctuation patterns, and simple engagement statistics, combined with conventional classifiers (e.g., SVMs, logistic regression, and random forests). While some studies considered temporal variations in textual features, these approaches largely relied on manual feature design and provided limited contextual modeling [1], [2].

With the rise of deep learning, later methods moved toward direct content encoding using convolutional and recurrent neural networks. CNN models are capable of capturing local patterns in short text, while RNNs and

LSTMs model sequential dependencies [3], often augmented by attention mechanisms to highlight words related to authenticity. These models reduce the effort of manual feature extraction and work efficiently when there is enough labeling data. However, due to their primari focus on textual content, they fail to articulate the viral dynamics and user behavior — factors that play a central role in the spread of rumors on social media.

B. PROPAGATION-BASED AND SPATIO-TEMPORAL MODELS

To overcome the limitations of content-only models, a second research direction integrated the propagation structure and information over time. Ma et al. (IJCAI 2016) approached rumor detection as a chain-layering problem [4], in which social contextual features are gradually synthesized over time and modeled using regression neural networks, enabling early detection even when only part of the propagation process is observed. Building on this formulation, Song et al. further explored credibility-aware early detection [5], explicitly optimizing model performance under limited visibility of the propagation process. Another prominent direction represents rumor events as propagation trees or graphs. Early work employed recursive neural networks [6] to model parent–child relations in reply trees, allowing both bottom-up and top-down information flow. More recently, graph neural networks such as BiGCN [7] have been applied to propagation structures, capturing directional diffusion patterns and yielding consistent improvements over flat sequence-based approaches. Subsequent studies extend this paradigm by incorporating edge directions, timestamps, or continuous-time diffusion dynamics to better model structural and temporal dependencies [8]. Despite their effectiveness, many propagation-based methods treat users as anonymous nodes and do not explicitly model user attributes or interaction behaviors. Moreover, most existing studies focus on microblog rumor detection, with limited investigation into whether propagation-based modeling can generalize to document-level fact-checking or news credibility assessment.

C. MULTIVIEW, USER-AWARE, AND KNOWLEDGE-ENHANCED MODELS

A subsequent research direction incorporates multiview modeling and explicit user information. Some methods construct propagation graphs and user interaction graphs simultaneously, allowing event representations to capture both information diffusion patterns and user behavioral characteristics. Architectures such as User-Reply-GCN follow this paradigm, while multiview fusion frameworks integrate textual content, propagation structure, and user features through attention- or gating-based mechanisms [9]. UMLARD is a representative example that shows that systematic multiview fusion improves robustness across datasets. Knowledge-enhanced methods further extend this line of work by incorporating external evidence or structured knowledge. Dual dynamic graph models, such as DDGCN [10] and the multiview spatio-temporal framework MST [11], represent each event using multiple evolving graphs, typically combining propagation-centric and user- or knowledge-centric views, and have demonstrated strong performance in early rumor detection.

Inspired by these approaches, MTF-GCN adopts a simplified temporal fusion strategy and is designed to support both rumor detection and news credibility assessment.

III. PROBLEM FORMULATION AND OVERVIEW OF MTF-GCN

A. PROBLEM DEFINITION

We consider a set of social media events $E = \{e_1, \dots, e_n\}$, where each event corresponds to a discussion thread initiated by a root post and followed by subsequent responses. For an event e_i, the root post and its replies are characterized by their identifiers, authors, textual content, and timestamps. All posts associated with the same event form a propagation tree that reflects how information spreads across the platform.

We focus on two closely related tasks. The first is rumor detection, where each event is labeled as either a rumor or a non-rumor. The second is news credibility assessment, in each news article and its associated social context, is labelled as reliable or unreliable. Although both problems use binary labels, the semantic meanings of these labels are different. The goal is to learn a classifier that maps each event to its corresponding label by simultaneously exploiting three sources of information: the text content, the structure and timing of propagation, and the characteristics and user interactions, to achieve good generalization to unseen events.

B. OVERVIEW OF MTF-GCN

The MTF-GCN represents each event using two mutually complementary dynamic graphs and clearly models their progression over time. The model framework comprises three stages: feature extraction, multi-perspective time-graph learning, and event-level layering. In the feature extraction stage, posts are encoded into text embeddings, while users are represented by dense vectors derived from profile attributes, forming node features for a propagation graph and a user interaction graph. For each event, both graphs are constructed dynamically and segmented into early, middle, and late diffusion stages. Graph convolutional layers are applied to each temporal snapshot to learn view-specific representations, which are then aggregated by a lightweight temporal fusion module. Finally, the fused propagation and user representations are concatenated and passed through fully connected layers to produce classification scores. By jointly modeling content, propagation dynamics, and user behavior over time, MTF-GCN provides a unified framework applicable to both rumor detection and document-level news credibility assessment.

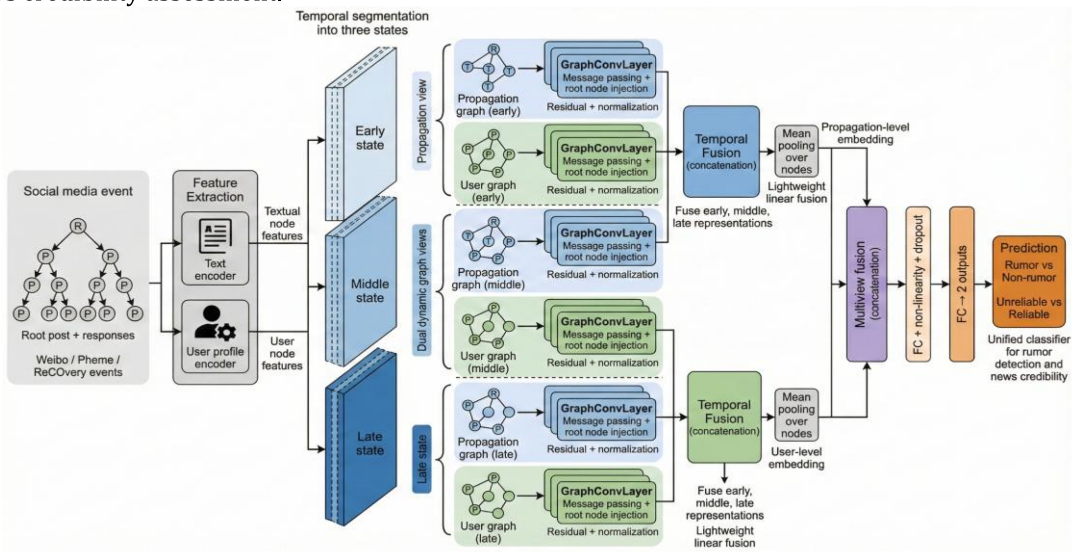

Figure 1: Overall framework of the proposed MTF-GCN model

IV. FEATURE REPRESENTATION

A. TEXTUAL FEATURE EXTRACTION

For each event, the textual content of posts provides the primary semantic signal. The root post is encoded with a transformer-based language model that maps the input text to a dense sentence-level embedding. This representation is then projected into a lower-dimensional space to facilitate integration with other graph features. Response posts are encoded similarly when text is available, using the same preprocessing pipeline. For extremely short or missing content, a fallback representation, such as the root embedding or a default vector, is used to ensure feature consistency across nodes. Optionally, response representations are enriched with averaged word-level embeddings to better capture local lexical cues in short messages. As a result, every post in the event is represented by a compact semantic vector, which constitutes the textual component of node features used in the propagation view of the model.

B. USER FEATURE EXTRACTION

Beyond textual content, user profiles provide useful signals about an account's tendency to spread reliable or unreliable information. For each user involved in an event, we extract a compact set of numerical and binary attributes that summarize social connectivity, account status, and basic profile characteristics, such as account age and activity indicators. Since these attributes vary in scale and distribution, simple preprocessing is applied. Count-based features are log-scaled to reduce the influence of outliers, and all features are normalized before being concatenated into a single user feature vector. This vector is then linearly projected into the same latent space as the textual embeddings. The resulting user representations capture general signals of behavior and status and are used to characterize nodes in user interaction graphs.

C. COMBINED NODE INITIALIZATION

The dynamic graphs in MTF-GCN use these two types of characteristics in a mutually complementary way. In the propagation graph, each node corresponds to a post in the event and is initialized using semantic embedding obtained from the text encoder. This representation emphasizes the connection between the messages as well as the change of content as the discussion progresses over time. In the user graph, nodes are aligned with the same set of posts, but the features attached to them come from the projected user profiles, reflecting who is participating rather than what is being said. For the root post, we keep its semantic representation available not only as the node feature in the propagation graph but also as a separate global descriptor of the event. This vector can later be injected into deeper layers of the network, for example by concatenation with intermediate node states, so that the influence of the original claim remains accessible throughout the graph convolution process. By initializing nodes in this way, MTF-GCN starts from a rich joint representation in which textual content and user characteristics are clearly separated yet easily combinable during subsequent multiview temporal learning.

V. DYNAMIC GRAPH CONSTRUCTION

A. TEMPORAL SEGMENTATION INTO STATES

An event is not only a static tree of posts but also a process that evolves over time. To capture this temporal dimension, we divide each event into three consecutive states based on its posting time. The duration of the event is determined from the original post to the final response, which is then divided evenly into three segments: the early stage (from the original post to one-third of the duration), the intermediate stage (from one-third to two-thirds of the duration), and the late stage (from two-thirds to the end).

For each time period, we construct a corresponding sub-event that includes only posts that appear during that period. The early state reflects the initial reactions to the information given, the intermediate state captures how the discussion develops as the event gains attention, and the late state represents the end of the exchange, which may include corrections, confirmation or deterioration of interest levels. This segmentation creates three "snapshots" of the same event, each differing in the posts and links triggered at each stage.

B. PROPAGATION GRAPH CONSTRUCTION

In the propagation view, each event is represented as a graph, where the nodes correspond to the posts in the discussion stream, including the original post and all responses. The edges encode the reply and re-share relationships between posts, and are treated as scalar edges. Self-loops are added to all nodes so that each post can preserve its own information during graph convolution. This propagation graph is instantiated across the three temporal states defined earlier. For each state, only posts whose timestamps fall within the corresponding time interval, along with the edges between them, are retained, yielding three propagation graphs for the early, middle, and late stages of the event. To enable efficient batch training, very large events are truncated and smaller ones are padded to a fixed maximum number of nodes at the indexing stage, while preserving the temporal structure. As a result, the model operates on a dynamic propagation graph evolving over three discrete snapshots, while remaining compatible with standard graph neural network operations.

C. USER INTERACTION GRAPH CONSTRUCTION

In parallel with the viral perspective, we build a second dynamic graph that focuses on user information. The user interaction graph uses the same node structure as the propagation graph, so each node still corresponds to a post in the event, but is tied to a user embedding that is inferred from the author's profile attributes. The edges in the user graph are inherited from reply and re-share relationships between posts, reflecting the interaction between the respective authors.

The same three-phase time segmentation is applied, creating user graphs for the early, intermediate, and late stages. By aligning the node set and time segment between the two views, the model maintains a common structural framework, and allows the propagation graph and user graph to encode complementary information sources to each other, thereby supporting multi-perspective time representation learning.

VI. MULTIVIEW TEMPORAL GRAPH LEARNING WITH MTF-GCN

A. DUAL GRAPH CONVOLUTIONAL LAYERS

After constructing the dynamic propagation and user graphs for each temporal state, MTF-GCN applies graph convolutional layers to update node representations in both views. For each state, we use a two-layer graph convolution module with non-linear activations, residual connections, and explicit access to the root post representationIn the propagation graph, the nodes are initialized using the semantic features obtained from the text encoder, whereas in the user graph, they carry profile-based features. The first graph convolution synthesizes information from neighboring nodes to create intermediate representations. The root post representation is then broadcast to all nodes and concatenated with these intermediate features, allowing each node to combine local neighborhood information with a global event context. A second convolution is then applied to the concatenated features, followed by a residual mapping to the projected input. Leaky ReLU and dropout are used for non-linearity and regularization, with batch normalization optionally applied to stabilize training.

This module is applied identically to both the propagation and user graphs at each temporal segment, yielding refined node representations that capture local structural patterns and a global, root-anchored view of the event.

B. TEMPORAL FUSION VIA CONCATENATION

For each viewpoint, the previous stage produces three episodes of node representations corresponding to early, intermediate, and late states. The MTF-GCN combines these representations into a single representation that summarizes the progression over time of the event. Instead of using a complex cross-attention mechanism, we adopt a simple time-fused module based on concatenation and linear transformations.

Specifically, the representations in each time state are first projected through separate linear layers to return to the same comparable space. The resulting embedding is then joined in the characteristic direction and passed through a final linear layer, followed by the activation function. This approach creates a compact unified performance that both captures the differences between time periods and is easy to optimize during the training process.

Compared to the multi-head attention mechanism, this concatenation-based fusion method has a lighter structure, uses fewer parameters, and avoids the fine-tuning of attention-specific hyperparameters. Despite its simplicity, the method is expressive enough to capture the critical temporal dynamics needed to distinguish fast-escalating events from those that develop more slowly or peak at a later stage.

C. DUAL-VIEW FUSION AND CLASSIFICATION HEAD

After consolidation over time, each event is represented by two fixed-size vectors, summarizing the evolution over time of the propagation graph and the user interaction graph. For each view, mean pooling on nodes is applied to achieve event-level embedding, which reduces local noise while preserving global patterns of structure and activity.

The propagation and user-level embedding are then concatenated and fed into a compact feed-forward layer of two fully connected layers. The first layer performs nonlinear transformation in combination with dropout, allowing interaction between features from two perspectives. The second layer maps the hidden representation to two output scores, which are then converted into predicted probabilities via a softmax function.

The same layer head is shared for different math tasks. For Weibo [1] and Pheme [6], the outputs correspond to rumored and non-rumored events, while for ReCOVery [12], they represent unreliable and credible news. Because the architecture does not depend on the specific semantics of the label, MTF-GCN can be easily applied to many different datasets, thereby demonstrating the flexibility of a unified model framework for misinformation detection.

VII. EXPERIMENTAL SETUP

A. DATASETS

We evaluate MTF-GCN on three datasets covering both rumor detection and news reliability assessments. Weibo [1], a Chinese microblog benchmark, contains 4,664 events (2,313 rumors and 2,351 non-rumors), each consisting of a source post and its replies and reposts, with about 800 posts per event on average. Pheme [6] is an English Twitter dataset of breaking-news discussions, including 297 rumor threads with 4,519 tweets in total (around 15 tweets per event), and is used in a standard binary rumor detection setting. ReCOVery [12] focuses on COVID-19 news credibility and comprises 2,029 English news articles labeled as reliable or unreliable, together with 140,820 related tweets; in this dataset, each article is treated as an event, making it a document-level fact-checking task.

Table 1. *Summary of datasets used in the experiments*

Dataset	**Domain**	**Language**	**Task**	**Instances**	**Positive class**	**Avg. context size**
Weibo [1]	Microblog	Chinese	Rumor detection	4,664	Rumor	≈ 816 posts
Pheme [6]	Microblog	English	Rumor detection	297*	Rumor	≈ 15 tweets
ReCOVery [12]	News + Social	English	News credibility	2,029	Unreliable	≈ 70 tweets

*Number of English rumor threads in the Pheme rumour scheme; non-rumor threads are added following standard preprocessing.

B. IMPLEMENTATION DETAILS

For Weibo and Pheme, source posts and replies are parsed from the original JSON files, normalized, and grouped by event. For ReCOVery, each news article is linked to its associated tweets via URLs and treated as a single event. Posts are ordered chronologically, with the number of nodes per event capped by truncating large events and padding smaller ones while preserving temporal order. Missing user metadata is filled with default values.

Each dataset is split into training, validation, and test sets using a 70/10/20 ratio. The model is trained with Adam, a small learning rate, and weight decay for up to 20 epochs, with early stopping based on validation performance. Batch size is set to a few hundred events. A shared configuration is used across experiments, with node embeddings projected into a moderate-dimensional space and two graph convolutional layers applied in both propagation and user views. Dropout is used for regularization, and the text encoder is fine-tuned jointly with the graph layers. Key hyperparameters are summarized in Table 2.

Table 2. *Main hyperparameters of MTF-GCN*

Hyperparameter	**Value**
Text / user embedding size	128
Hidden size in GCN layers	128
Number of GCN layers	3
Maximum nodes per event	100
Dropout rate	0.2
Optimizer	Adam
Learning rate	5×10^{-5}
Weight decay	1×10^{-4}
Batch size	256
Maximum epochs	20
Early stopping patience	5
Train / Val / Test split	0.7 / 0.1 / 0.2

VIII. RESULTS AND ANALYSIS

A. OVERALL PERFORMANCE ON RUMOR DETECTION

As shown in Table 3, MTF-GCN consistently outperforms DDGCN [10] on both Weibo and Pheme, with clear and meaningful performance gaps. On Weibo, the improvement is moderate (around 2–3 percentage points), indicating a stable gain even on a relatively well-structured dataset. Precision and recall remain closely aligned, and F1 scores for the rumor and non-rumor classes are nearly symmetrical, suggesting minimal class bias. On Pheme, the margin becomes substantially larger, with approximately 10 percentage points gains in both accuracy and macro-F1 compared to DDGCN. This wider gap highlights MTF-GCN's greater robustness under noisier, more heterogeneous conditions. Although a slight difference appears between rumor and non-rumor F1 on Pheme, the imbalance remains limited and does not affect overall stability. The consistent gains across both datasets, especially the larger improvement on Pheme, confirm the effectiveness of jointly modeling propagation dynamics and user interactions within a unified temporal graph framework for rumor detection.

Table 3. *Overall performance on Weibo and Pheme*

Dataset	**Model**	**Accuracy**	**F1**	**Pre**	**Recall**	**F1 Non-rumor**	**F1 Rumor**
Weibo	DDGCN [10]	0.948	0.950	0.953	0.948	–	–
	MTF-GCN (ours)	0.9775	0.9775	0.9775	0.9775	0.9775	0.9730
Pheme	DDGCN [10]	0.855	0.844	0.846	0.841	–	–
	MTF-GCN (ours)	0.9522	0.9485	0.9474	0.9496	0.9623	0.9348

B. PERFORMANCE ON FACT-CHECKING

On the ReCOVery dataset, MTF-GCN demonstrates strong generalization in document-level verification. The model achieves over 93% accuracy and above 92% macro-F1, demonstrating stable and competitive overall performance. A class-wise analysis shows a clear but meaningful asymmetry. The Unreliable class reaches an F1 of around 95%, noticeably higher than the Reliable class (approximately 89%). This gap reflects a conservative prediction bias, in which the model prioritizes detecting unreliable content. Such behavior is desirable in early-warning or risk-sensitive scenarios, since overlooking unreliable information is typically more costly than issuing false alarms.

Compared to SAFE [12], whose macro-F1 remains around 75%, MTF-GCN yields a substantial improvement - nearly 17 percentage points in macro-F1 - along with consistent gains across both classes. This margin indicates that integrating propagation dynamics and user-interaction modeling provides a clear advantage for document-level information verification, positioning MTF-GCN as an effective filtering component in practical fact-checking systems.

Table 4. *Overall performance on ReCOVery dataset*

Model	Reliable F1	Unreliable F1	macro-F1	Accuracy
SAFE [12]	0.833	0.672	0.753	-
MTF-GCN (Ours)	0.8889	0.9500	0.9209	0.9310

IX. DISCUSSION

C. IMPACT OF TEMPORAL STATES AND GRAPH SIZE

MTF-GCN models the progression of events through three temporal states - early, intermediate, and late. Combining all posts into a single static graph makes it impossible for the model to distinguish between events with similar structures but different propagation rates, which in turn leads to poorer performance, especially in early detection. Using only two states yields limited improvement, whereas finer segmentation increases computational cost and yields sparse, low-information graphs. Thus, the three-time states achieve an effective balance between temporal resolution and model robustness.

A similar trade-off applies to graph size. Events that are too large incur high computational costs and complicate batching. By limiting the number of buttons per event and selecting posts in chronological order, the model preserves representative context while maintaining a stable training process. In fact, moderate graph size limits capture most of the benefits of graph-based modeling, whereas larger graphs yield only gradual improvements.

D. PRACTICAL CONSIDERATIONS AND APPLICATIONS

Beyond performance on the evaluation metrics, MTF-GCN is designed for practical implementation. By working on incomplete discussions and explicitly modeling the early spread phase, the model supports early warning of rumors on social media platforms, allowing it to flag potentially harmful events before they reach their peak spread. In assessing the reliability of news, MTF-GCN assigns each article a trust score based on the social context and the behavior of users who share it. This score can be used to prioritize items that need manual verification, where high recall for untrusted content helps ensure that the majority of problematic cases are reviewed.

The model also adds value by identifying abnormal or coordinated propagation patterns, as it intelligently integrates profile information and user interactions. From a systems perspective, MTF-GCN still ensures computational efficiency, with lightweight graph components and predictable memory usage, making the model suitable for large-scale, real-time, or batch monitoring pipelines.

X. CONCLUSION AND FUTURE WORK

In this work, we introduce MTF-GCN, a unified framework for rumor detection and news credibility assessment on social media. Each event is modeled by two complementary dynamic graphs that capture message propagation and user interactions, along with a set of temporal states. Graph convolutional layers operate across views and time, and a lightweight temporal fusion module summarizes event evolution. Experiments on Weibo, Pheme, and ReCOVery show that MTF-GCN consistently achieves strong performance, matching or outperforming competitive baselines and generalizing well to document-level credibility assessment, with high recall for unreliable content.

Future work may explore integrating external knowledge and explicit fact-checking evidence, improving interpretability by identifying influential users and key posts, and extending the framework to support multi-class veracity labels for real-world misinformation monitoring.

XI. ACKNOWLEDGMENTS

This research was funded by Ho Chi Minh City University of Foreign Languages — Information Technology under grant number H2024-14.

XII. REFERENCES

[1] A. Zubiaga, M. Liakata, and R. Procter (2017), "Exploiting Context for Rumour Detection in Social Media," pp. 105–109, doi: 10.1007/978-3-319.

[2] A. Zubiaga, M. Liakata, R. Procter, G. Wong Sak Hoi, and P. Tolmie (Mar. 2016), "Analysing how people orient to and spread rumours in social media by looking at conversational threads," *PLoS One*, vol. 11, doi: 10.1371/journal.pone.0150989.

[3] J. P. Singh, A. Kumar, N. P. Rana, and Y. K. Dwivedi (Apr. 2022), "Attention-Based LSTM Network for Rumor Veracity Estimation of Tweets," *Information Systems Frontiers*, vol. 24, pp. 459–474, doi: 10.1007/s10796-020-10040-5.

[4] J. Ma *et al. (*2016), "Detecting Rumors from Microblogs with Recurrent Neural Networks," in *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, New York, USA: AAAI Press, pp. 3818–3824. doi: 10.5555/3061053.3061153.

[5] C. Song, C. Yang, H. Chen, C. Tu, Z. Liu, and M. Sun (Aug. 2021), "CED: Credible Early Detection of Social Media Rumors," *IEEE Trans. Knowl. Data Eng.*, vol. 33, pp. 3035–3047, doi: 10.1109/TKDE.2019.2961675.

[6] J. Ma, W. Gao, and K.-F. Wong (Jul. 2018), "Rumor Detection on Twitter with Tree-structured Recursive Neural Networks," Iryna Gurevych and Yusuke Miyao, Eds., Association for Computational Linguistics, pp. 1980–1989. doi: 10.18653/v1/P18-1184.

[7] T. Bian *et al. (*2020), "Rumor detection on social media with bi-directional graph convolutional networks," in *AAAI 2020 - 34th AAAI Conference on Artificial Intelligence*, AAAI press, pp. 549–556. doi: 10.1609/aaai.v34i01.5393.

[8] Q. Huang, C. Zhou, J. Wu, L. Liu, and B. Wang (Jun. 2023), "Deep spatial–temporal structure learning for rumor detection on Twitter," *Neural Comput. Appl.*, vol. 35, pp. 12995–13005, doi: 10.1007/s00521-020-05236-4.

[9] X. Chen, F. Zhou, G. Trajcevski, and M. Bonsangue (Mar. 2022), "Multi-view learning with distinguishable feature fusion for rumor detection," *Knowl. Based. Syst.*, vol. 240, doi: 10.1016/j.knosys.2021.108085.

[10] Sun, M., Zhang, X., Zheng, J., & Ma, G. (2022). DDGCN: Dual Dynamic Graph Convolutional Networks for Rumor Detection on Social Media. Proceedings of the AAAI Conference on Artificial Intelligence, 36(4), 4611-4619.

[11] X. Huang *et al.* (Oct. 2025), "Multiview Spatio-Temporal Learning With Dual Dynamic Graph Convolutional Networks for Rumor Detection," *IEEE Trans. Comput. Soc. Syst.*, vol. 12, pp. 2469–2479, doi: 10.1109/TCSS.2024.3520105.

[12] X. Zhou, A. Mulay, E. Ferrara, and R. Zafarani (Oct. 2020), "ReCOVery: A Multimodal Repository for COVID-19 News Credibility Research," in *International Conference on Information and Knowledge Management, Proceedings*, Association for Computing Machinery, pp. 3205–3212. doi: 10.1145/3340531.3412880.

MTF-GCN: MẠNG TÍCH CHẬP ĐỒ THỊ HỢP NHẤT THEO THỜI GIAN ĐA GÓC NHÌN CHO PHÁT HIỆN TIN ĐỒN VÀ KIỂM CHỨNG THÔNG TIN TRÊN MẠNG XÃ HỘI

Tiếu Phùng Mai Sương, Nguyễn Thị Thúy A

TÓM TẮT —Tin đồn và các nguồn thông tin có độ tin cậy thấp lan truyền rất nhanh trên mạng xã hội và có thể gây ra những tác động nghiêm trọng đến xã hội. Nhiều phương pháp hiện có phụ thuộc chủ yếu vào nội dung văn bản hoặc cấu trúc lan truyền tĩnh, do đó còn hạn chế trong việc nắm bắt sự tương tác động giữa thông tin và người dùng theo thời gian. Trong nghiên cứu này, chúng tôi đề xuất MTF-GCN, một mô hình hợp nhất theo thời gian đa góc nhìn cho bài toán phát hiện tin đồn và kiểm chứng thông tin ở mức tài liệu. Mỗi sự kiện được biểu diễn bằng hai đồ thị động: (i) đồ thị lan truyền được xây dựng từ quan hệ trả lời và chia sẻ lại, và (ii) đồ thị người dùng dựa trên đặc điểm hồ sơ và tương tác. Quá trình lan truyền được chia thành ba giai đoạn thời gian gồm sớm, trung gian và muộn; sau đó được xử lý bằng các lớp tích chập đồ thị và tổng hợp thông qua một cơ chế hợp nhất theo thời gian gọn nhẹ. Các biểu diễn thu được từ nhánh lan truyền và nhánh người dùng tiếp tục được kết hợp để thực hiện phân loại nhị phân. Thực nghiệm trên các bộ dữ liệu Weibo, Pheme và ReCOVery cho thấy MTF-GCN đạt độ chính xác cao (lần lượt 97,75%, 95,22% và 93,10%), đặc biệt mạnh trong việc phát hiện nội dung không đáng tin cậy. Những kết quả này cho thấy việc mô hình hóa không gian–thời gian đa góc nhìn kết hợp với cơ chế hợp nhất theo thời gian đơn giản là hiệu quả và phù hợp cho các hệ thống cảnh báo sớm và giám sát thông tin sai lệch.

Từ khóa — Phát hiện tin đồn, kiểm chứng thông tin, mạng xã hội, mạng nơ-ron đồ thị, mô hình hóa không gian–thời gian đa góc nhìn, tương tác người dùng.

ThS. Tiếu Phùng Mai Sương nhận bằng Thạc sĩ Khoa học máy tính tại trường Đại học Khoa học tự nhiên TP. Hồ Chí Minh vào năm 2017. Hiện nay, cô là giảng viên tại Trường Đại học Ngoại ngữ - Tin học TP. Hồ Chí Minh (HUFLIT).

Hướng nghiên cứu chính: Xử lý ngôn ngữ tự nhiên, Trí tuệ nhân tạo.

ThS. Nguyễn Thị Thúy A nhận bằng thạc sĩ ngành Khoa học máy tính tại trường ĐH Khoa học tự nhiên, ĐHQG-HCM vào năm 2018. Hiện tại Thạc sĩ Thúy A đang là giảng viên tại khoa Công nghệ thông tin, trường ĐH Ngoại ngữ - Tin học TP. Hồ Chí Minh (HUFLIT).

Hướng nghiên cứu chính: Xử lý ngôn ngữ tự nhiên, Trí tuệ nhân tạo.