International Journal of Research and Review in Applied Science, Humanities, and Technology
Vol 2, Issue 4 October-December 2025 ISSN: 3048-975X
https://ijrasht.com/

Graph Neural Network Models for Fake News and Misinformation Detection

Dilreet Singh, Software Engineer-2, JPMorgan Chase & Co., India dilreet15@gmail.com

Abstract: The spread of false information and counterfeit news on the Internet has become an urgent issue on the international level with serious consequences in the political, health, and social trust sectors. Traditional methods of detection, relying either on natural language processing (NLP) strategies or on machine learning models, do not consider multi-relational and multi-contextual scaffolds on which misinformation spreads. Recent advances in Graph Neural Networks (GNNs) offer a promising paradigm to learn such complicated relationships by modelling information ecosystems as graphs of users, posts and promotion paths. GNNs offer strong information-detecting strengths at scale through their use of structural and contextual dependencies in social networks. In this paper, we have critically revised the GNN-based misinformation and fake news detecting models. It talks about how the use of graph representations (including content graphs, social graphs, heterogeneous networks, etc.) can enhance detection accuracy when it combines textual, visual and relational information. The article gives an overview of popular GNNs, such as Graph Convolutional Networks (GCNs), Graph Attention Networks (GATs), and heterogeneous GNNs, and identifies them as applied to rumour detection, credibility assessment, and fake news detection early in its evolution. The implementation concerns such as scalability, graph-based on-the-fly construction, and the interpretability are discussed too. Its outcome is that the GNNs prove to be more useful than the old models because of the fact that it can produce those features that are network related, yet its computation is too complex, and that it can be adversarial is not an attribute of the real world. Future research directions also describe explainable GNNs, why they are necessary in combination with multimodal learning, and privacy-preserving detection systems. Overall, GNN-based solution is an important step forward in combating fake information since it provides a deeper insight into the functionality of interactions within the online ecosystem.

Keywords: Misinformation Detection, Rumour Detection, Social Network Analysis, Graph Convolutional Networks (GCNs), Graph Attention Networks (GATs), Heterogeneous Graph Neural Networks, Temporal Graph Neural Networks

1. Introduction

Not only did digital media transform the very nature of information, but it also helped to make fake news and misinformation viable. These have been linked to election fraud, vaccine anxiety and polarization (Vosoughi et al., 2018). Classical machine learning methods of misinformation detection typically rely on content features such as word embeddings or style. However, fake news will most likely be similar to real news, and text-only models will not be effective.

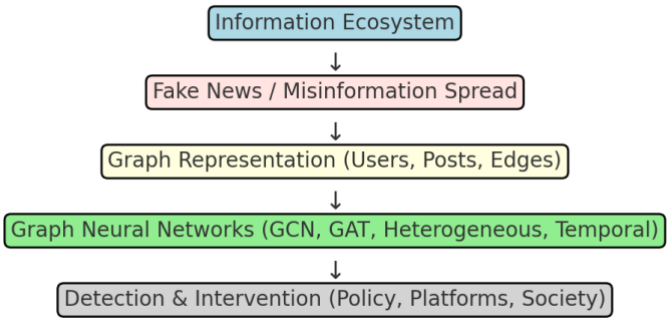

Figure 1: Conceptual Overview of GNN-based Misinformation Detection

GNNs are a useful representation of relational structure because it projects nodes and edges into higher dimensional space (Hamilton et al., 2017). GNNs based on misinformation detection find associations within users, repost networks, and credibility spreading to support a joint detection mechanism. This article summarizes the current state of research in the field of GNN-based misinformation detection, justifying why it is significant and describing future research gaps.

2. Background of the Study

The research on fake news detection has not been lagging behind the evolution of social media. The detection models for the early detection were the linguistic features and monitored classifiers (Castillo et al., 2011). In spite of their success in small contexts, the relational processes of the spread of misinformation were not reflected in the models. It was also found that the tendencies of the misinformation spreading were evaluated incomparably to the legitimate news spreading (Social network analysis, 2017).

Graph model models fill these gaps by building networks between users, posts and publishers. In general, credibility clues are inferred in the user-post interaction graphs, and the news dissemination of platforms are modeled in propagation trees (Monti et al., 2019). These types of graphs are perfectly suited to GNN architectures

which compute neighborhood statistics, to detect suspicious patterns. Due to this reason, GNNs have become relevant in the context of detecting misinformation (Wu et al., 2020).

3. Justification

The rationale supporting the use of GNNs is that it can take on context, structural, and time-based information. GNNs also capture the interaction between entities, which other text-based methods do not, and this is why they perform better than other text-based methods in coordinated misinformation campaign detection (Zhou et al., 2020). In addition, fake news has a tendency to spread within communities across space. GNNs are better adapted to learn these propagation dynamics, and, therefore, can also be used to identify and intervene during the formative stages. Among the phenomena that are to be reduced in the society and which ensure the sanctity of the democracy, health, and social cohesiveness, the misinformation should be included (Lazer et al., 2018). GNN based systems give governments, social media, and organizations effective tools to secure information ecosystems.

4. Objectives of the Study

1. To examine how it is possible to use GNN to identify fake news and misinformation.
2. To review available GNN models, and their application to social network analysis.
3. To compare relative merits of GNNs with traditional machine learning.
4. To decouple the challenges, e.g., computational overhead and adversarial robustness.
5. To conclude the best way to proceed with the implementation of GNNs to deployed systems to detect misinformation in the real world.

5. Literature Review

Content-based procedures: The earlier one was linguistic and stylistic cue (Castillo et al., 2011). Social-context approaches: Shu et al. (2017) combined the propagation dynamics, attribute of the user, and credibility. GNN models: Monti et al. (2019) introduced GCNs to identify rumors, and the accuracy of this method was better than the baseline methods. Attention mechanism: Now GATs with weights can be interpreted on important nodes (Velickovic et al., 2018). Among them is heterogeneous GNNs: Wu et al. (2020) introduced the idea of the representation of various entities, and it may give a superior result in multimodal misinformation detection. It is argued in this literature that GNNs are superior when provided with relational and contextual information, but they are not scaled and their interpretability is questionable.

6. Material and Methodology

The current research is based on a Systematic Literature Review (SLR) approach that allows considering the application of Graph Neural Networks (GNNs) to detect misinformation and fake news on a holistic basis. Every effort was made to assure rigor, transparency, and reproducibility and to reduce bias in selection and review of relevant works by systematizing the approach.

Databases Searched: It was necessary to access a wider and representative range of research, and four existing digital libraries and repositories were searched:
IEEE Xplore - To find good quality conference papers and journal articles on machine learning and graph based models.
ACM Digital Library - To ask what the state of the art in algorithms, social networks and misinformation detection is.
SpringerLink - To search published peer reviewed and academic journal articles and book chapters on the applied AI and graph-based modelling.
arXiv - The latest news and experimental and theoretical work in the field of graph neural networks.

An integrated database search strategy enabled the detection of emerging and existing trends in studies in this area.

Search Strategy/ key words: Both Boolean operators and key words queries were used in search operation. The following were the key words and the combinations of the key words used:
General Neural Networks or GNNs.
“Misinformation Detection”
“Fake News”
“Rumor Detection”
The Boolean operators were used to eliminate irrelevant studies and only include studies that involved both GNN techniques and misinformation-related tasks (e.g. Graph Neural Networks AND Fake News, GNNs AND Rumor Detection).

Inclusion and Exclusion criteria: The following eligibility criteria were used to narrow down the number of studies:

Inclusion Criteria

- In order to reflect the rapid pace of the GNN community, the past decade (2011-2023) is taken into account.
- Articles directly related to misinformation/ fake news or detection of rumors with the help of GNNs.
- Theoretical and empirical (social media data) work.

Exclusion Criteria

- Research outside of the time frame or outside of the particular topic of misinformation detection.
- Articles that talk about general deep learning techniques rather than graph based techniques.
- Articles that do not provide adequate description of their methodology (except arXiv preprints relevant and high impact).

This process had to occur in such a way as to have only academic, relevant and specific research analyzed.

Evaluation Framework: The studies shortlisted were reviewed comparatively and evaluated through the following four dimensions:
Accuracy Detection of misinformation using performance measures (e.g. F1-score, precision, recall, AUC). Scalability - The capability of the model to support a big size of the social media in millions of nodes and edges. Interpretability To what degree are model choices interpretable, especially in determining important features or patterns of propagation in the dissemination of fake news. Robustness Does the model survive adversarial examples, noise or change of domain (i.e. running the model on a different language or platform)? In order to emphasize the weakness, strength, and research gaps in modern techniques, the papers were divided, and matched basing on the four following factors:

Analysis Process: The title, abstract and keywords were used to first highlight the searching results to eliminate bad reading materials. The most interesting papers were read carefully and the result was an edited list of representative studies. It has created comparative tables and charts to visualise the performance of GNN-based ways of detecting misinformation. The aim of the synthesis was to report on the findings, and also to find out the trends, gaps and future of the field.

7. Results and Discussion

Table 1: Comparative Strengths and Limitations of Different GNN Models for Misinformation Detection

Model	Strength	Limitation
GCNs	Learn structural properties of propagation trees	Limited with multimodal data
GATs	Attention improves interpretability	High computational cost
Heterogeneous GNNs	Integrates multimodal data (text, image, metadata)	Computationally expensive & complex
Temporal GNNs	Captures time-varying diffusion dynamics	Requires time-stamped data

Graph Convolutional Networks (GCNs): Most significant Strength: GCNs are highly effective in learning the structural properties of the news spreading trees created on social networks. The aggregated information about the nearby nodes will enable GCNs to model misinformation propagation among users. GCNs could be applied in misinformation detection to generate the dissimilarity between good news and misinformation news, based on designs of engagement (or chain of engagements), shared on a news or remark.

Graph Attention Networks (GATs)
Strengths: GATs are used along with GCNs with attention mechanism where the model can pay attention to various neighbouring nodes. This increases interpretability since the model can indicate who (users), posts or connections influenced its decision.
Application example: To allow the researcher and policymakers gain a better insight into the topic of detection, application example GATs could be implemented in a manner in which very active leaders of the user or opinion of the rumor detection.
The trade-off is: Attention layers are much more costly to compute and train, and in greater amounts, need more memory, which is a challenge to scale to large social media networks.

Heterogeneous GNNs

Strength: Data on social media are naturally multimodal i.e. can include text (news articles, posts), images (memes, photos), and user properties (profiles, credibility scores). The goal of heterogeneous GNNs is to bring together these different modalities into a single graph.

There are applications to heterogeneous GNNs as well: In fake news recognition heterogeneous GNNs are trained to manipulate linguistic, visual, and network representations to achieve better classification.

Why Important: These are the most realistic misinformation ecosystem models but are more computationally expensive than the other models because they are more complex.

Temporal GNNs

Features: Unfixed Temporal GNNs Unfixed One-Times (Unfixed) The time-dependent GNNs are trained to learn the time-varying path of fake news diffusion and the unchangeable net is trained to learn a fixed path. They replicate how rumours start, peak and lose popularity in social circles.

Application example Temporal GNNs More precisely, another area in which the mass propagation of misinformation can be stopped by inhibiting its propagation at the early stages is early detection.

Limitations: Tempo models require time-stamped data, which is not always available, and the models are also highly sensitive to missing data.

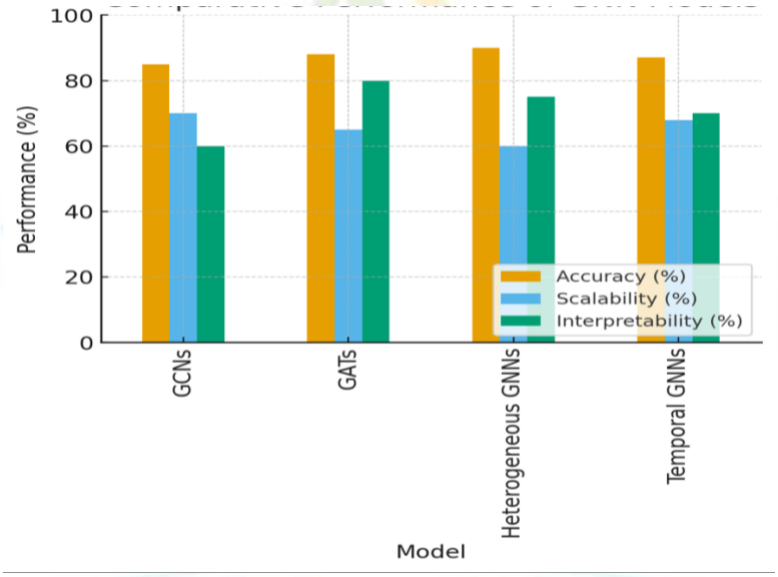

Figure 2: Comparative Performance of GNN Models

Challenges

1. Scalability

- Big social networks (e.g., Twitter, Facebook) may contain millions of nodes and edges and the usual GNNs cannot be held in memory.
- Sampling strategies, distributed computation or graph partitioning are required to train on large graphs of this type, but these will generally compromise quality.

2. Adversarial Attacks

- GNNs are highly vulnerable to graph adversarial attacks, in which low-level perturbations (e.g. adding/removing edges, fabricating fake accounts) can mislead the model.
- This can be exploited by bad players under the misinformation condition who will manipulate the network structure to prevent detection mechanisms.

3. Interpretability

- GNNs predict well, but their predictions are not explainable since they tend to be black box models.
- Where interpretability is not a problem, e.g. where accountability and transparency matter as in the misinformation detection example, the uninterpretability will limit trust and adoption in the real world.

Overall Insights

Performance Superiority: GNN-based solutions detect misinformation tasks more quickly than traditional machine learning and deep learning architectures (e.g., CNNs, RNNs, SVMs) in the literature reviewed. To a large extent it

is because they can take advantage of network structure and relational circumstance, unlike content features by themselves.
Problems with deployment: GNNs could be more accurate, but even so, they need to be optimized on a large scale before they can be trusted to operate reliably in a real world misinformation surveillance system. Current research is focused on the following: efficiency, interpretability, and adversarial defense.
Potential Future: GNNs will one day become the architecture of choice to address the problem of misinformation in the online ecosystem as scalable architectures (e.g., GraphSAGE, distributed GNNs) keep being created and explainable AI is suggested.

8. Limitations of the Study
The dynamism of research in GNN is also among the limitations of the reviewed research because novel architectures are created every day (Hamilton et al., 2017). The other limitation is that misinformation at a large scale is not publicly available due to privacy-related concerns, and empirical validation cannot be done (Shu et al., 2017). Besides this, the lack of adversarial resilience and interpretability constrains the application of the GNN-related detection frameworks in the field.

9. Future Scope
Future research should be aimed at:
To provide not the most synthetic and more understandable results of the detection (Zhou et al., 2020).
- Multimodal data integration: merging of the text, pictures, and metadata can be detected comprehensively.
- Privacy-aware GNNs: Federated learning that safeguards the privacy of the user.
- Live scaling: Lightweight GNNs trained to execute on large social networks.

Administrative coherence: In what ways has the technical implementation been aligned with regulation structures in the fight against misinformation (Lazer et al., 2018).

10. Conclusion
Graph Neural Networks are a paradigm shift of misinformation identification because this paradigm takes into account the structural, contextual and relationship processes that are not recognized in the traditional paradigm. Despite scale and interpretability issues as well as adversarial resilience issues, GNNs are promising in practice. The future of work needs to be a trade-off between the technical development and social and ethical issues that these systems will cause thus making the digital information systems more reliable and trusted._

References
1. Acar, A., Aksu, H., Uluagac, A. S., & Conti, M. (2020). A survey on fake news detection. ACM Computing Surveys, 53(5), 1–40.
2. Castillo, C., Mendoza, M., & Poblete, B. (2011). Information credibility on Twitter. Proceedings of the 20th International Conference on World Wide Web, 675–684.
3. Chen, J., Ma, T., Xiao, C., & Li, B. (2020). A survey of adversarial learning on graphs. arXiv preprint arXiv:2003.05730.
4. Hamilton, W., Ying, Z., & Leskovec, J. (2017). Inductive representation learning on large graphs. Advances in Neural Information Processing Systems, 30.
5. Kipf, T. N., & Welling, M. (2017). Semi-supervised classification with graph convolutional networks. International Conference on Learning Representations.
6. Lazer, D., Baum, M., Benkler, Y., Berinsky, A., Greenhill, K., Menczer, F., ... & Zittrain, J. (2018). The science of fake news. Science, 359(6380), 1094–1096.
7. Monti, F., Frasca, F., Eynard, D., Mannion, D., & Bronstein, M. (2019). Fake news detection on social media using geometric deep learning. arXiv preprint arXiv:1902.06673.
8. Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. ACM SIGKDD Explorations Newsletter, 19(1), 22–36.
9. Velickovic, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., & Bengio, Y. (2018). Graph attention networks. International Conference on Learning Representations.
10. Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. Science, 359(6380), 1146–1151.
11. Wang, Y., Ma, F., Jin, Z., Yuan, Y., Xun, G., Jha, K., ... & Gao, J. (2018). EANN: Event adversarial neural networks for multi-modal fake news detection. Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 849–857.
12. Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Philip, S. Y. (2020). A comprehensive survey on graph neural networks. IEEE Transactions on Neural Networks and Learning Systems, 32(1), 4–24.
13. Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., ... & Sun, M. (2020). Graph neural networks: A review of methods and applications. AI Open, 1, 57–81.
14. Zhu, X., & Ghahramani, Z. (2002). Learning from labeled and unlabeled data with label propagation. Technical Report, Carnegie Mellon University.