Dynamic Text-Attributed Graphs and Learning Models for Community Evolution: A Survey of Recent Advances

Sruthi K S[1], P B Divya[2], A Sreekumar[1], Kannan Balakrishnan[1]
Cochin University of Science and Engineering, Kochi, Kerala, India[1]
Union Christian College, Aluva, Kerala, India[2]
Corresponding author: Sruthi K S, Email: sruthyksreedharan@gmail.com

Understanding community changes in dynamic networks is a key research area with applications in social networks, academic networks, biological networks, and financial networks. While traditional graph models focus on their structural patterns, real-world systems often combine these patterns with substantial textual content associated with nodes and edges, resulting in Dynamic Text-Attributed Graphs (DyTAGs). Such networks record both time-based interactions and the development of textual details, allowing for a more thorough analysis of how communities are created, changed, and evolved. This work reviews recent progress in dynamic graph learning for community detection in text-attributed networks. This paper discusses the framework for handling Dynamic Text-Attributed Graphs (DyTAGs). It surveys advanced models, including TGAT, DyGFormer, and methods involving Large Language Models (LLMs). The research emphasizes the importance of integrating temporal, structural, and semantic data and shows that this multimodal integration improves both the precision and the understanding of the community evolution of networks.

Keywords: Complex Networks, Dynamic Text Attributed Graphs, Community Evolution, Graph Neural Networks, Large Language Models.

2025. In Kusum Kumari Bharti & Noor Firdoos Jahan (eds.), *Next-Gen Data Analytics and Intelligent Automation*, 48–62. Computing & Intelligent Systems, SCRS, India. DOI: https://doi.org/10.56155/978-81-975670-6-3-5

1 Introduction

Complex network analysis is a growing research area because many real-world systems, such as biological ecosystems, technological infrastructures, and social communities, can be represented as networks. This analysis helps us understand both the structure and functional dynamics of these systems. These networks, with their complex structures and emerging patterns, show a strong framework for modeling interactions between connected entities. A critical development in this area is Text-Attributed Networks [6], where each node has both links to other nodes and associated text that adds meaningful context.

Text-attributed networks naturally emerge across various domains. In social media, users are connected through interactions while also producing text such as posts, comments, and profile information. In scientific publishing, papers form citation or co-authorship networks, enriched with textual details like titles, abstracts, and keywords. In e-commerce, products are linked by co-purchase patterns and described through reviews or specifications. These networks combine structural complexity with semantic depth, making them well-suited for exploring community structures and their evolution over time.

In these networks, communities [30]-tightly connected groups of nodes with similar characteristics—act as key structural components. Detecting these communities is a central task in network science, helping to uncover patterns such as groups of researchers working in related areas or users with common interests. In dynamic environments, however, communities are not fixed; they constantly evolve by growing, shrinking, merging, splitting, or even disappearing. Understanding how communities change over time is essential for identifying trends, tracking the spread of influence, and modeling the life cycle of group behavior.

This study aims to advance the field by presenting a comprehensive survey focused on detecting and analyzing community evolution in text-attributed complex networks. By combining temporal network modeling with semantic text representation, we seek to uncover subtle patterns in how communities form and change over time. We evaluate our approach using real-world datasets, including scholarly citation networks and online discussion forums, showcasing its ability to reveal dynamic, interpretable, and semantically meaningful community trajectories.

2 Preliminaries

This section presents the core concepts and notation necessary for modeling and understanding Dynamic Text-Attributed Graphs (Dy-TAGs) [23] [24]. Dy-TAGs are a type of evolving graph in which nodes and edges carry both temporal interaction data and rich textual information. They are especially well-suited for capturing the complex interplay of structure, time, and content in systems like social media, academic networks, and communication platforms.

2.1 Text Attributed Networks (TAGs)

A *Text-Attributed Dynamic Network* is defined as a time-evolving attributed graph:

$$\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{T}, \mathcal{X}_v, \mathcal{X}_e)$$

Where:

- $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$ is the set of **nodes**, representing entities such as users, documents, or products.
- $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of **edges**, representing interactions or relationships (e.g., citations, communications).
- $\mathcal{T} = \{t_1, t_2, \ldots, t_T\}$ is the set of **discrete time steps** capturing network dynamics.
- $\mathcal{X}_v : \mathcal{V} \times \mathcal{T} \to \mathbb{D}_v$ is the **node text attribute function**, where for each node v_i at time t, $\mathcal{X}_v(v_i, t) = d_i^{(t)} \in \mathbb{D}_v$, and $\mathbb{D}_v$ denotes the textual data associated with nodes.
- $\mathcal{X}_e : \mathcal{E} \times \mathcal{T} \to \mathbb{D}_e$ is the **edge text attribute function**, where for each edge $e_{ij} = (v_i, v_j)$ at time t, $\mathcal{X}_e(e_{ij}, t) = d_{ij}^{(t)} \in \mathbb{D}_e$, and $\mathbb{D}_e$ represents the textual data associated with edges.

To enable numerical modeling and learning on text-attributed networks, it is essential to convert unstructured textual data into a structured, machine-readable format. This is achieved by mapping each piece of text to a dense vector representation in a continuous vector space using an appropriate *text embedding function.*

Let $f_{\text{text}} : \mathbb{D} \to \mathbb{R}^d$ denote a generic embedding function, where:

- $\mathbb{D}$ is the domain of unstructured text (e.g., user posts, article abstracts, or messages), and
- $\mathbb{R}^d$ is a d-dimensional semantic vector space.

This function takes a textual input and produces a real-valued vector that captures its semantic features. Specifically:

$$
\begin{aligned}
\text{Node embedding:} \quad & \mathbf{x}_i^{(t)} = f_{\text{text}}(d_i^{(t)}), \quad \mathbf{x}_i^{(t)} \in \mathbb{R}^{d_v}, \\
\text{where } d_i^{(t)} &= \mathcal{X}_v(v_i, t) \\
\text{Edge embedding:} \quad & \mathbf{x}_{ij}^{(t)} = f_{\text{text}}(d_{ij}^{(t)}), \quad \mathbf{x}_{ij}^{(t)} \in \mathbb{R}^{d_e}, \\
\text{where } d_{ij}^{(t)} &= \mathcal{X}_e(e_{ij}, t)
\end{aligned}
$$

2.2 Community Evolution

In many real-world networks, such as social media platforms, research collaborations, and online discussion forums, entities like users, articles, or posts often form communities. These communities are groups of nodes that are more densely connected to each other than to the rest of the network, typically reflecting meaningful relationships such as shared interests, similar research topics, or common involvement in specific discussions. Identifying such communities provides valuable insights into the structure and behavior of complex systems.

Incorporating textual information from nodes and edges greatly improves both the accuracy and interpretability of community detection [29]. Two nodes that are only loosely connected structurally may still belong to the same community if their textual content is semantically similar. On the other hand, tightly connected groups might reflect different themes if their associated text differs significantly. By integrating structural connections with semantic similarity, we can identify communities that are not only well-connected but also consistent in content, resulting in more meaningful and contextually relevant groupings.

2.3 Dynamic Graph Neural Networks (Dy-GNNs)

A *Dynamic Graph Neural Network (Dynamic GNN)* [21] [22] is a model designed to learn from graph-structured data that evolves. Unlike static GNNs that operate on a fixed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, dynamic GNNs are applied to time-evolving graphs where both structure and features change over time.

2.3.1 Discrete-Time Dynamic Graphs

Let the dynamic graph be represented as a sequence of T graph snapshots:

$$\{\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_T\}, \quad \text{where } \mathcal{G}_t = (\mathcal{V}_t, \mathcal{E}_t, \mathbf{X}_t)$$

- $\mathcal{V}_t$: Set of nodes at time t
- $\mathcal{E}_t \subseteq \mathcal{V}_t \times \mathcal{V}_t$: Set of edges at time t
- $\mathbf{X}_t \in \mathbb{R}^{|\mathcal{V}_t| \times d}$: Node feature matrix at time t

A dynamic GNN layer computes node embeddings at time t as:

$$\mathbf{H}_t = \text{GNN}_\theta(\mathbf{H}_{t-1}, \mathcal{G}_t) = \sigma\left(\hat{\mathbf{A}}_t \mathbf{H}_{t-1} \mathbf{W}_t\right)$$

Where:

- $\hat{\mathbf{A}}_t$: Normalized adjacency matrix at time t
- $\mathbf{H}_{t-1}$: Node embeddings from the previous timestep
- $\mathbf{W}_t$: Learnable weight matrix
- σ: Non-linear activation function (e.g., ReLU)

Temporal dependencies are often modeled using recurrent units:

$$\mathbf{h}_i^{(t)} = \text{GRU}\left(\mathbf{h}_i^{(t-1)}, \mathbf{z}_i^{(t)}\right)$$

where $\mathbf{z}_i^{(t)}$ is the embedding derived from GNN at time t.

2.3.2 Continuous-Time Dynamic Graphs

In a continuous-time setting, the graph evolves via timestamped events:

$$\mathcal{G}(t) = (\mathcal{V}(t), \mathcal{E}(t), \mathbf{X}(t))$$

Events at times $\{t_1, t_2, \ldots, t_k\}$ trigger updates in the graph. Upon an event affecting node i at time t, its representation is updated as:

$$\mathbf{h}_i(t^+) = \text{GRU}\left(\mathbf{h}_i(t^-), \phi(\mathbf{m}_{i,j}, \Delta t)\right)$$

Where:

- $\mathbf{h}_i(t^-)$, $\mathbf{h}_i(t^+)$: Node embeddings before and after time t
- $\mathbf{m}_{i,j}$: Message from node j to i
- $\Delta t = t - t_{\text{last}}$: Time since last update
- ϕ: Message function (e.g., an MLP)

2.4 Pretrained Embedding Methods and Large Language Models (LLMs)

Modern large language models (LLMs) have revolutionized the way we represent text. They generate detailed, multi-dimensional embeddings that effectively capture the intricate meanings between words, sentences, and entire documents. When working with text-attributed networks, these embeddings are crucial. They transform the unstructured text associated with nodes or edges into a structured vector format, making it compatible with graph-based models, such as Graph Neural Networks (GNNs). This integration allows for a more comprehensive analysis, leveraging both the textual content and the structural relationships within the network. Table 1 refers to the important state-of-the-art language embedding models.

Table 1: Comparison of Embedding Models

Model	Dimension	Highlights
BERT-Base [17]	768	General-purpose transformer; requires pooling for sentence embeddings; not tuned for similarity
BERT-Large [17]	1024	Higher representational power; captures deeper semantics; more computationally intensive
SBERT [18]	768	Siamese architecture; optimized for sentence-level similarity and clustering
DistilBERT [19]	768	Lightweight and efficient; faster inference with reasonable semantic quality
E5-Base [20]	768	Balanced, trained with contrastive objectives; strong semantic grouping and clustering ability
`text-embedding-3-large` [16]	3072	High semantic precision; excellent for classification, search, and clustering
`text-embedding-ada-002` [16]	1536	Highly scalable, cost-effective embedding for general-purpose tasks

Table 2: Summary of Key Contributions on Dy-TAGs

Authors	Paper	Year	Contribution
Xu et al. [1]	Unlocking Multi-Modal Potentials for Dynamic Text-Attributed Graph Representation	2025	Proposed MoMent, a multi-modal dynamic graph framework for Dy-TAGs with modality-specific encoders and alignment loss.
Lei, Runlin et al. [3]	Exploring the Potential of Large Language Models as Predictors in Dynamic Text-Attributed Graphs	2025	Proposed Graph Agent-Dynamic (GAD), a multi-agent LLM framework for Dy-TAGs that performs prediction without task-specific training.
Roy et al. [7]	LLM-driven Knowledge Distillation for Dynamic Text-Attributed Graphs	2025	Proposed LKD4DyTAG: distills LLM knowledge into lightweight, temporally-aware GNNs.
T Yu, Jianxiang, et al. [8]	Leveraging large language models for node generation in few-shot learning on text-attributed graphs.	2025	LLM4NG: few-shot learning with node generation via LLMs.
Zhang, Siwei, et al. [9]	Unifying Text Semantics and Graph Structures for Temporal Text-attributed Graphs with Large Language Models	2025	CROSS: LLM model for capturing temporal text dynamics in TTAGs.
Li, Shujie, et al. [10]	HeTGB: AComprehensive Benchmark for Heterophilic Text-Attributed Graphs	2025	HeTGB benchmark: heterophilic TAGs across real-world domains.
Xu, Yuanyuan, et al. [11]	UniDyG: A Unified and Effective Representation Learning Approach for Large Dynamic Graphs	2025	The paper proposes UniDyG, a unified and scalable dynamic graph representation learning model that handles both Continuous-Time and Discrete-Time Dynamic Graphs by introducing a novel Fourier Graph Attention (FGAT) mechanism.
Wu, Xixi, et al. [15]	When Do LLMs Help With Node Classification?	2025	A Comprehensive Analysis LLMNodeBed: Many experiments to evaluate LLMs in node classification.
Liang, Xun, et al. [4]	Controlled Text Generation for Large Language Model with Dynamic Attribute Graphs	2024	Introduced DATG for controlled text generation using dynamic attribute graphs.
Zhang, Delvin Ce, et al. [5]	Text-Attributed Graph Representation Learning: Methods, Applications, and Challenges	2024	Surveyed TAGRL methods, applications, challenges, and fusion strategies.
Zhang et al. [2]	DTGB: A Comprehensive Benchmark for Dynamic Text-Attributed Graphs	2024	Introduced DTGB, the first benchmark for Dy-TAGs with datasets, tasks, and model evaluations.
Li, Zhuofeng, et al. [12]	TEG-DB: AComprehensive Dataset and Benchmark of Textual-Edge Graphs	2024	TEG-DB benchmark textual-edge graphs with node/edge texts.
Wang, Yaoke, et al. [13]	Bridging Local Details and Global Context in Text-Attributed Graphs	2024	GraphBridge: Bridges local/global semantics in TAGs.
Pan, Bo, et al. [14]	GraphNarrator: Generating Textual Explanations for Graph Neural Networks	2024	TAGExplainer: interpretable framework for TAGs using LLMs.
Yan, Hao, et al. [6]	A Comprehensive Study on Text-attributed Graphs: Benchmarking and Rethinking	2023	Introduced the CS-TAG benchmark for static TAGs, comparing PLMs and GNNs on eight datasets; evaluates structure vs. text-based learning; public code.

3 Literature of Techniques for Community Evolution in TAGs

The integration of textual and structural information in graph representation learning has seen significant advancements, particularly with the emergence of Text-Attributed Graphs (TAGs) and their dynamic variants (DyTAGs). A considerable challenge in this domain lies in effectively capturing the interplay between evolving graph structures and rich semantic text associated with nodes and edges.

[1] identify critical shortcomings in existing dynamic TAG methods that encode text using pretrained language models (PLMs) combined with dynamic graph models in an edge-centric manner. These approaches often neglect long-range temporal and semantic dependencies. To address this, they propose MoMent, a model that explicitly incorporates three core modalities of DyTAGs—temporal, textual, and structural—via specialized encoders. MoMent demonstrates significant performance gains over strong baselines such as TGAT and DyGFormer on tasks including link prediction and edge classification across seven real-world datasets.

Benchmarking efforts have been pivotal to advancing this research area. [2] introduce DTGB, the first large-scale benchmark tailored for dynamic text-attributed graphs. DTGB provides a suite of diverse datasets and standardized evaluation tasks—such as future link prediction and textual relation generation—that require integrated modeling of temporal, structural, and semantic signals.

Exploring the potential of LLMs, [3] propose Graph Agent Dynamic (GAD), a multi-agent LLM-based system designed for DyTAGs. GAD addresses challenges like limited context windows and domain heterogeneity through collaborative agents that summarize knowledge and dynamically adapt.

In the realm of controlled text generation conditioned on graph attributes, [4] propose DATG. This novel framework guides large language models (LLMs) using Dynamic Attribute Graphs to influence output semantics.

Text-attributed graph representation learning (TAGRL) has been comprehensively surveyed by [5], who categorize methods into early fusion, late fusion, and joint learning strategies depending on the integration approach of structural and textual data. They discuss applications spanning node classification, community detection, and recommendation systems, while identifying key challenges including noise in text, heterogeneity alignment, and scalability.

Addressing the lack of standardized evaluation, Yan, Hao, et al. (2023) [6] present CS-TAG, a benchmark framework for TAGs integrating PLMs and GNNs.

Their work systematically assesses various learning paradigms—including PLM-based, GNN-based, co-training, and topological pretraining—across eight large-scale TAG datasets.

On the front of knowledge distillation, [7] introduce LKD4DyTAG, which distills semantic knowledge from a pretrained LLM (teacher) into a lightweight, temporally aware GNN (student) for dynamic text-attributed graphs. Their method aligns textual edge embeddings from the LLM with spatio-temporal representations from the GNN via an efficient time-encoding scheme.

Focusing on few-shot learning, [8] propose LLM4NG. This plug-and-play framework leverages large language models (LLMs) to generate class-specific exemplar nodes via prompt-based text generation. These synthetic nodes are embedded using sentence encoders, such as Sentence-BERT, and linked to the original graph via an edge predictor trained on the raw structure.

To unify text semantics with evolving graph structures, [9] present CROSS, which models Temporal Text-Attributed Graphs (TTAGs) by combining a Temporal Semantics Extractor—using LLM prompts to capture semantic evolution—and a Semantic-Structural Co-encoder that fuses textual and structural features via a cross-modal mixer.

Recognizing the challenges in heterophilic graphs where linked nodes are dissimilar, [10] introduce HeTGB, a benchmark for heterophilic text-attributed graphs characterized by low homophily and rich textual content. HeTGB includes five real-world datasets and evaluates GNNs, PLMs, and co-training methods for node classification.

UniDyG ([11]) provides a unified and scalable approach to representation learning on large, dynamic graphs, with a focus on efficient temporal modeling. Meanwhile, TEG-DB ([12]) provides a comprehensive benchmark for Textual-Edge Graphs (TEGs), highlighting the importance of textual data on edges in addition to nodes.

GraphBridge ([13]) proposes a multi-granularity framework for TAGs that bridges local textual semantics and global structural context. It features a graph-aware token reduction module and a multi-step integration pipeline that combines language models and Graph Neural Networks (GNNs).

For explainability, TAGExplainer ([14]) introduces a novel framework that generates natural language explanations for TAG learning models. By extracting saliency scores on nodes, edges, and tokens, and leveraging large language models (LLMs) for pseudo-label explanations refined through an expert iteration strategy, TAGExplainer produces faithful, concise, and human-readable justifications that outperform other LLM baselines across multiple datasets.

Finally, [15] conduct an extensive analysis of the utility of LLMs for node classification on text-attributed graphs. Their benchmark, LLMNodeBed, evaluates 14 datasets and multiple LLM paradigms under various learning settings. Results highlight that LLM-based methods excel especially in semi-supervised scenarios, and that Explainer paradigms perform best when textual information is crucial.

4 Benchmark Datasets

To enable a thorough and consistent assessment of models on Dynamic Text-Attributed Graphs (Dy-TAGs), we utilize a collection of diverse benchmark datasets drawn from various real-world contexts. The DTGB benchmark [28] offers a unified framework that standardizes data preprocessing, evaluation metrics, and task formulations, ensuring fair comparisons and reproducibility among different dynamic graph learning approaches. The selected datasets facilitate a broad spectrum of tasks, including link prediction, edge classification, and node clustering, which collectively demonstrate the diverse challenges inherent in Dy-TAG research.

Textual Edge Graph (TEG) [12] benchmark datasets comprise real-world dynamic networks, where each edge represents an interaction annotated with a timestamp and detailed textual information.

Table 3: Summary of DTGB Benchmark Datasets for Dynamic Text-Attributed Graphs

Dataset	Domain	Text Source	Supported Tasks
Reddit-Post	Social Media	Post content	Link prediction, edge classification
Reddit-Comment	Social Media	Comment threads	Community detection, reply analysis
Wikipedia	Knowledge Graph	Article edits	Node classification
DBLP	Academic Network	Paper abstracts	Co-authorship prediction
ACL-Anthology	Academic Network	Titles, abstracts, venue info	Temporal node clustering
GitHub	Software Development	Commit messages	Link prediction
Amazon-QA	E-commerce	Questions and answers	Node classification, QA relevance
StackOverflow	Developer Forum	Q&A threads	Expert routing, temporal prediction

5 Challenges and Future Directions

A core challenge in modeling Dynamic Text-Attributed Graphs (Dy-TAGs) is effectively integrating heterogeneous modalities, including structural, temporal, and textual data. Many current approaches consider textual and temporal information as secondary features, resulting in incomplete or fragmented representations of the data. Developing unified and coherent embeddings that capture both the evolving semantic content and graph topology is particularly difficult, especially when the textual attributes are sparse or contain noise.

Scalability remains a significant challenge in handling Dy-TAGs, which are inherently high-dimensional and large-scale, particularly in real-time scenarios like social media or citation networks. Managing the dynamic interactions among millions of nodes and edges, each potentially associated with timestamped textual information, demands substantial computational power and memory resources. Many current GNN-based and transformer architectures struggle to deliver scalable performance without sacrificing temporal precision or the richness of semantic information.

A key direction for future research is the development of multi-modal learning frameworks that equally prioritize text, temporal, and structural information. Rather than simply combining features at a late stage, upcoming models should adopt early fusion or co-training strategies, enabling each modality to influence the learning of representations mutually. This approach might involve jointly optimizing temporal transformers alongside large language models (LLMs) within graph neural network architectures.

Lastly, improving explainability and generalization is essential for real-world applications. Future models should incorporate causal inference and explanation mechanisms to enhance transparency in tasks such as link prediction and node classification. Additionally, robust pretraining and transfer learning strategies will be vital to ensure Dy-TAG models can adapt and generalize across different domains, supporting diverse applications from recommendation engines to scientific knowledge tracking.

6 Conclusion

This study offers an in-depth exploration of Dynamic Text-Attributed Graphs (Dy-TAGs), which represent evolving networks enriched with temporal and textual data. We investigated key challenges in effectively learning from these graphs, including the fusion of multiple modalities, the alignment of temporal information, and the capture of semantic changes over time. Through an analysis of prominent models—including TGAT, DyGFormer, GraphMixer, and CNEN—and evaluation frameworks like DTGB and MoMent, we highlighted both advancements and persistent gaps in current methodologies.

Looking forward, we advocate for the development of unified architectures that are explicitly modality-aware, greater incorporation of large language models, and scalable learning techniques as essential for progress in Dy-TAG research. These innovations have the potential not only to boost predictive accuracy but also to deepen our understanding of dynamic processes in practical domains such as social media, academic collaborations, and digital ecosystems.

References

[1] Xu, Y., Zhang, W., Zhang, Y., Lin, X., and Xu, X. (2025). Unlocking Multi-Modal Potentials for Dynamic Text-Attributed Graph Representation. *arXiv preprint arXiv:2502.19651.*

[2] Zhang, J., et al. (2024). DTGB: A Comprehensive Benchmark for Dynamic Text-Attributed Graphs. *arXiv preprint arXiv:2406.12072.*

[3] Lei, R., et al. (2025). Exploring the Potential of Large Language Models as Predictors in Dynamic Text-Attributed Graphs. *arXiv preprint arXiv:2503.03258.*

[4] Liang, X., et al. (2024). Controlled text generation for large language model with dynamic attribute graphs. *arXiv preprint arXiv:2402.11218.*

[5] Zhang, D. C., et al. (2024). Text-attributed graph representation learning: Methods, applications, and challenges. *Companion Proceedings of the ACM Web Conference 2024.*

[6] Yan, H., Wang, T., Liu, S., Tang, J., and Yu, P. S. (2023). A comprehensive study on text-attributed graphs: Benchmarking and rethinking. *Advances in Neural Information Processing Systems*, 36:17238–17264.

[7] Roy, A., Yan, N., and Mortazavi, M. (2025). LLM-driven knowledge distillation for dynamic text-attributed graphs. *arXiv preprint arXiv:2502.10914.*

[8] Yu, J., He, H., Du, X., and Wang, H. (2025). Leveraging large language models for node generation in few-shot learning on text-attributed graphs. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 39, No. 12.

[9] Zhang, S., Hu, Z., Guo, S., and Li, Y. (2025). Unifying text semantics and graph structures for temporal text-attributed graphs with large language models. *arXiv preprint arXiv:2503.14411.*

[10] Li, S., Zhang, M., Zhang, Y., and Sun, L. (2025). HeTGB: A Comprehensive Benchmark for Heterophilic Text-Attributed Graphs. *arXiv preprint arXiv:2503.04822*.

[11] Xu, Y., Zhang, H., Li, T., and Zhou, J. (2025). UniDyG: A Unified and Effective Representation Learning Approach for Large Dynamic Graphs. *IEEE Transactions on Knowledge and Data Engineering*.

[12] Li, Z., Wang, Q., Zhou, L., and Gao, H. (2024). TEG-DB: A comprehensive dataset and benchmark of textual-edge graphs. *Advances in Neural Information Processing Systems*, 37:60980–60998.

[13] Wang, Y., Zhao, K., Feng, X., and Liu, J. (2024). Bridging local details and global context in text-attributed graphs. *arXiv preprint arXiv:2406.12608*.

[14] Pan, B., Li, J., Wang, X., and Zhou, Y. (2024). GraphNarrator: Generating Textual Explanations for Graph Neural Networks. *arXiv preprint arXiv:2410.15268*.

[15] Wu, X., Lin, C., Zhang, W., and Shen, H. (2025). When Do LLMs Help With Node Classification? A Comprehensive Analysis. *arXiv preprint arXiv:2502.00829*.

[16] Radford, A., Narasimhan, K., Salimans, T., and Sutskever, I. (2018). Improving language understanding by generative pre-training. *OpenAI Technical Report*, page 3.

[17] Devlin, J., Chang, M. W., Lee, K., and Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Volume 1 (Long and Short Papers), pages 4171–4186.

[18] Reimers, N. and Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using Siamese BERT-networks. *arXiv preprint arXiv:1908.10084*.

[19] Sanh, V., Debut, L., Chaumond, J., and Wolf, T. (2019). DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. *arXiv preprint arXiv:1910.01108*.

[20] Wang, L., Liu, Y., Zhang, M., and Zhao, T. (2022). Text embeddings by weakly-supervised contrastive pre-training. *arXiv preprint arXiv:2212.03533*.

[21] Wang, Y., Zhang, Z., Liu, T., Wu, X., and Sun, H. (2024). Bridging local details and global context in text-attributed graphs. *arXiv preprint arXiv:2406.12608*.

[22] Zhu, X., Xue, H., Zhao, Z., Xu, W., Huang, J., Guo, M., Wang, Q., Zhou, K., and Zhang, Y. (2025). LLM as GNN: Graph vocabulary learning for text-attributed graph foundation models. *arXiv preprint arXiv:2503.03313*.

[23] Zhang, Z., Liu, Y., Sun, Y., and Zhang, Y. (2023). Out-of-distribution generalized dynamic graph neural network with disentangled intervention and invariance promotion. *arXiv preprint arXiv:2311.14255*.

[24] Wang, Z., Li, Y., Jiang, Q., Liu, Y., and Li, Z. (2024). Towards Dynamic Graph Neural Networks with Provably High-Order Expressive Power. *arXiv preprint arXiv:2410.01367*.

[25] Veesam, S. B., Chatterjee, A., Mondal, A., Mukhopadhyay, A., and Ghosh, A. (2025). Design of an integrated model with temporal graph attention and transformer-augmented RNNs for enhanced anomaly detection. *Scientific Reports*, 15(1), 2692.

[26] Yu, L., Wu, C., Zhou, P., Wang, Y., Li, H., Yu, P. S., and Zhao, W. (2023). Towards better dynamic graph learning: New architecture and unified library. *Advances in Neural Information Processing Systems*, 36, 67686–67700.

[27] Cheng, K., Ren, P., Zhang, Y., and Ma, J. (2024). Co-neighbor encoding schema: A light-cost structure encoding method for dynamic link prediction. *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*.

[28] Yi, L., et al. (2025). TGB-Seq Benchmark: Challenging Temporal GNNs with Complex Sequential Dynamics. *arXiv preprint*, arXiv:2502.02975.

[29] Sruthi, K. S., Sreekumar, A., and Balakrishnan, K. (2025). Text-attributed community detection in complex networks through LLMs and GNNs: A powerful fusion of language and graphs. *Neurocomputing*, Article ID: 130573.

[30] Girvan, M. and Newman, M. E. J. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12), 7821–7826.

[31] Zhou, J., Cui, G., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C., and Sun, M. (2020). Graph neural networks: A review of methods and applications. *AI Open*, 1, 57–81.

[32] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30.