

Dynamic Heterogeneous Graph Representation Learning: A Survey

Huan Liu¹, Pengfei Jiao^{1*}, Jie Yin², Hongjiang Chen¹ and Zhidong Zhao^{1,3}

¹School of Cyberspace, Hangzhou Dianzi University, China

²Discipline of Business Analytics, The University of Sydney, Australia

³Zhejiang Provincial Key Laboratory for Sensitive Data Security Protection and Confidentiality Management, Hangzhou, China

{huanliu, pjiao, hchen, zhaozd}@hdu.edu.cn, jie.yin@sydney.edu.au

Abstract

Graph representation learning (GRL) serves as a canonical paradigm for modeling complex networks. However, real-world AI systems inherently manifest as evolving heterogeneous entities with complex interactions, posing significant challenges to static or homogeneous modeling. To address these complexities, representation learning for Dynamic Heterogeneous Graphs (DHGs) has emerged as a vital approach for learning low-dimensional representations that simultaneously preserve structural semantics and temporal dynamics. This survey presents the first systematic review of DHG representation learning methods. We first introduce a unified formal definition that encompasses both discrete-time and continuous-time DHGs from the perspective of temporal granularity. Building upon this formulation, we propose a novel algorithm-centric taxonomy that categorizes existing literature, including early embedding-based approaches, graph neural network (GNN)-based models, and relatively recent Transformer-based DHG methods, while explicitly highlighting their intrinsic modeling biases with respect to dynamic granularity. Furthermore, we summarize representative applications of DHG representation learning, along with commonly used datasets and benchmarks. Finally, we discuss promising research directions that guide future advances in this rapidly evolving field.

1 Introduction

Graphs have been established as a foundational data structure for modeling complex real-world networks and have seen widespread applications across many fields in recent years, such as social network analysis [Li *et al.*, 2023c; Wang, 2025], foundation models [Fey *et al.*, 2024; Liu *et al.*, 2025b], and recommendation systems [Liu *et al.*, 2021; Li *et al.*, 2024c]. By preserving the complex structure and dependencies of graphs in a low-dimensional vector space, graph representation learning (GRL) has been a cornerstone for enabling generalizable reasoning over relational data in modern

intelligent systems such as large language models (LLMs), and has attracted considerable research interest [Fey *et al.*, 2024; Jin *et al.*, 2024; Liu *et al.*, 2025b].

However, real-world systems typically manifest as dynamically evolving, multi-typed entities with complex interactions [Hong *et al.*, 2020; Qi *et al.*, 2025], where dynamics and heterogeneity are inseparably coupled rather than independent properties [Xu *et al.*, 2022; Wang *et al.*, 2025b]. Thus, Dynamic Heterogeneous Graph (DHG) representation learning has emerged as a vital paradigm for capturing the joint interplay between heterogeneous semantics and temporal evolution [Fey *et al.*, 2024; Liu *et al.*, 2025b]. Nevertheless, existing research typically models heterogeneity and dynamics separately [Wang *et al.*, 2022a; Kazemi *et al.*, 2020], and effective DHG representation learning remains non-trivial due to several key challenges. First, heterogeneous semantics and temporal dynamics are intrinsically coupled; thus, methods that simply concatenate heterogeneous modeling with sequential models fail to capture relation-specific evolutionary behaviors. Second, as heterogeneous interactions follow different temporal behaviors, homogeneous modeling is insufficient to capture the sudden, periodic, and long-term cumulative dynamic changes of multi-typed nodes/edges. Third, discrete- and continuous-time DHGs exhibit distinct structural evolution patterns at different temporal granularities; while the former captures coarse-grained global updates, the latter tracks rapid, asynchronous, and localized changes. Consequently, designing unified architectures that preserve granularity-specific dynamics while effectively filtering temporal noise remains an open challenge. Finally, explicitly parameterizing relation-specific semantics and dynamics increases computational overhead, thus hindering the scalability of large-scale DHG learning.

Although a growing body of studies has addressed certain facets of these challenges [Dileo *et al.*, 2023; Wang, 2025], the literature remains fragmented and lacks a unified review. To bridge this gap, we present the first comprehensive survey dedicated to DHG representation learning. We establish a unified formal definition of DHGs that bridges discrete-time and continuous-time formulations. On this basis, we propose an algorithm-centric taxonomy of DHG representation learning, categorizing methods into embedding-based, GNN-based, and Transformer-based paradigms, and explicitly characterizing their inductive biases under different temporal granularities.

*Corresponding author.

We further discuss representative applications and summarize commonly used open-source datasets and benchmarks for DHG representation learning, revealing widespread inconsistencies in graph construction and fixed temporal partitions that hinder fair comparison and reliable performance evaluation.

Distinct from prior GRL surveys, this work provides a unified and systematic review of DHG representation learning. Existing surveys on heterogeneous graphs focus on heterogeneous modeling principles and representation techniques under static settings [Yang *et al.*, 2020a; Wang *et al.*, 2022a]. In parallel, surveys on dynamic GRL offer valuable insights into temporal modeling paradigms and evaluation protocols [Kazemi *et al.*, 2020; Barros *et al.*, 2021], but primarily under homogeneous assumptions. This work also diverges from benchmarking studies such as TGB 2.0 [Gastinger *et al.*, 2024], which emphasize evaluation frameworks rather than algorithmic mechanisms. While these existing efforts have significantly advanced their respective areas, the joint modeling of heterogeneity and temporal evolution remains under-explored, necessitating a dedicated and holistic review of DHG representation learning. Our major contributions include:

- We present a formal definition of DHG encompassing both discrete-time and continuous-time formulations, and introduce a structured taxonomy organizing methods into four families based on their core architecture. An overview is presented in Figure 1.
- We systematically review representative methods within each family, analyzing their technical mechanisms and identifying shared limitations that constrain current DHG learning. We further compile open-source datasets and benchmarks, exposing inconsistent construction and deterministic evaluation protocols that hinder fair comparison and reliable assessment.
- We identify three progressive research frontiers toward DHG foundation models: efficiency for enabling scaling law investigation, generalizability for cross-domain transfer and continual learning, and trustworthiness for enabling explainable DHG learning and causal discovery.

2 Preliminaries and Notations

This section introduces the formal definition for DHGs and the representation learning tasks.

Definition 1 (Dynamic Heterogeneous Graph). *A Dynamic Heterogeneous Graph (DHG) is defined as a graph jointly modeling structural heterogeneity and temporal evolution:*

$$\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A}, \mathcal{R}, \mathcal{T}, \phi, \varphi),$$

where $\mathcal{V}$ and $\mathcal{E}$ denote the (possibly time-varying) sets of nodes and edges over a temporal domain $\mathcal{T}$. The type-mapping functions $\phi : \mathcal{V} \rightarrow \mathcal{A}$ and $\varphi : \mathcal{E} \rightarrow \mathcal{R}$ assign each node and edge to a node-type set $\mathcal{A}$ and an edge-type set $\mathcal{R}$, respectively.

Remark. This formulation is general and subsumes multiple mainstream graph types as special cases:

Heterogeneous vs. Homogeneous. A graph $\mathcal{G}$ is *heterogeneous* if it involves multiple semantic types, i.e., $|\mathcal{A}| + |\mathcal{R}| > 2$. When $|\mathcal{A}| = 1$ and $|\mathcal{R}| = 1$, $\mathcal{G}$ degenerates to a *homogeneous graph* with a single node and edge type.

Discrete vs. Continuous. The temporal domain $\mathcal{T}$ governs the dynamic nature of the graph. If $|\mathcal{T}| = 1$, $\mathcal{G}$ reduces to a *static graph* $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A}, \mathcal{R}, \phi, \varphi)$ representing a single heterogeneous graph snapshot. If $\mathcal{T}$ is a discrete ordered set, then $\mathcal{G}$ is a *discrete-time DHG*, represented as a sequence of heterogeneous graph snapshots $\mathcal{G} = \{\mathcal{G}^{(t)}\}_{t \in \mathcal{T}}$. If $\mathcal{T}$ is continuous, the graph is a *continuous-time DHG* consisting of a sequence of timestamped edges:

$$\mathcal{E} = \{(u, v, r, t) \mid u, v \in \mathcal{V}, r \in \mathcal{R}, t \in \mathcal{T}\},$$

thereby capturing fine-grained, asynchronous interactions.

Definition 2 (DHG Representation Learning). *Given a DHG $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A}, \mathcal{R}, \mathcal{T}, \phi, \varphi)$ with optional node features $\mathbf{X}^{(t)}$, the goal of DHG representation learning is to learn an encoder $f_{\Theta} : \mathcal{G} \rightarrow \mathbb{R}^{|\mathcal{V}| \times d}$ that maps each node $v \in \mathcal{V}$ to a low-dimensional embedding $\mathbf{h}_v^{(t)} \in \mathbb{R}^d$, such that $\mathbf{h}_v^{(t)}$ jointly preserves heterogeneity and dynamics proximity, supporting various downstream tasks.*

3 DHG Representation Learning Taxonomy

DHG representation learning requires jointly encoding type-specific semantics and temporal dependencies, as different relation types exhibit distinct evolutionary patterns. Based on the algorithmic mechanism for capturing such coupled dynamics, we categorize methods into: (1) embedding-based approaches preserving structural and temporal proximity through random walks or matrix factorization; (2) GNN-based approaches with type-aware message-passing and aggregation across snapshots; (3) Transformer-based approaches unifying type and temporal signals within attention mechanisms; and (4) application-oriented approaches formulating domain problems as DHG learning tasks. Table 1 summarizes representative methods.

3.1 Embedding-based

Embedding-based approaches learn node representations by preserving structural proximity derived from neighborhoods. Based on temporal granularity, we distinguish: (1) random walk-based and (2) incremental update-based methods, both operating on discrete snapshots to capture co-occurrence patterns; and (3) temporal point process-based methods modeling continuous-time events and fine-grained interactions.

Random Walk-based

Random-walk-based methods leverage predefined meta-structure to capture the co-evolution of network topology and heterogeneous semantics across temporal snapshots. The core objective is to maximize the co-occurrence likelihood of nodes in sequences generated by constrained random walks:

$$\mathcal{L} = \sum_{v \in \mathcal{V}^{(t)}} \sum_{c \in \mathcal{N}_{\mathcal{S}}^{(t)}(v)} \log P(c \mid v; \theta), \quad (1)$$

where $\mathcal{S} \in \{\mathcal{P}, \mathcal{M}\}$ denotes the meta-paths or meta-graphs and $\mathcal{N}_{\mathcal{S}}^{(t)}(v)$ represents the time-aware context of node v induced by meta-structure $\mathcal{S}$ at snapshot t . Foundational approaches like Change2vec [Bian *et al.*, 2019] perform walks on historical-current snapshot pairs to preserve short-term

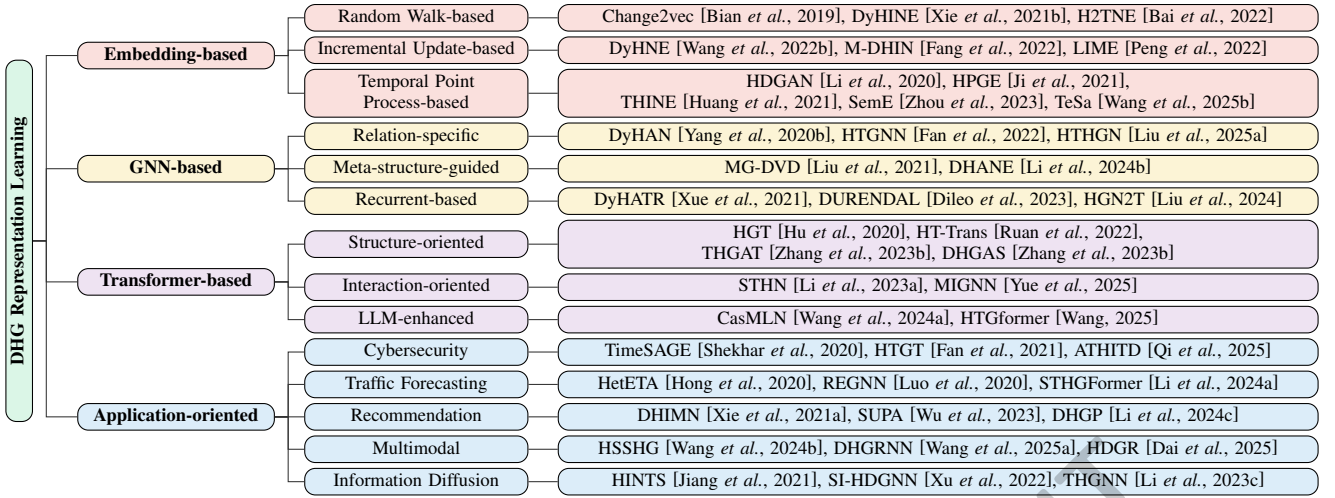


Figure 1: A Taxonomy of Dynamic Heterogeneous Graph (DHG) Representation Learning Methods.

temporal continuity. To capture complex dependencies, DyHINE [Xie *et al.*, 2021b] and M-DHIN [Fang *et al.*, 2022] introduce attention mechanisms and meta-graph guidance to weight relevant paths dynamically. Recent advances, such as H2TNE [Bai *et al.*, 2022], employ hyperbolic embeddings to capture hierarchical scale-free DHG structures.

While effective in preserving structural proximity, these methods rely on expert-defined meta-paths/graphs and often struggle to incorporate rich node attributes, limiting their expressiveness for complex semantic tasks.

Incremental Update-based

Incremental update frameworks efficiently model network evolution by focusing on localized calibration rather than re-training. The core of these methods is to combine the previous state with DHG perturbations to update node representations:

$$\mathbf{h}_u^{(t)} = f_{\text{upd}}(\mathbf{h}_u^{(t-1)}, \Delta\mathcal{G}^{(t)}). \quad (2)$$

DyHNE [Wang *et al.*, 2022b] models evolution as perturbations of meta-path augmented adjacency matrices, efficiently updating embeddings via generalized eigenvalue perturbation. LIME [Peng *et al.*, 2022] employs a recursive neural network to selectively adjust shared semantic embeddings within a local cuboid space. Besides, Change2vec [Bian *et al.*, 2019], DyHINE [Xie *et al.*, 2021b] and M-DHIN [Fang *et al.*, 2022] explicitly identify nodes affected by edge additions/deletions and node insertions/removals and restrict embedding updates to local neighborhoods rather than the entire network.

Despite their scalability advantages, these methods suffer from cumulative approximation drift over long temporal horizons and struggle with abrupt structural changes that violate temporal smoothness assumptions.

Temporal Point Process-based

Temporal point process (TPP) frameworks treat DHG evolution as a stream of asynchronous events, modeling the conditional intensity of heterogeneous interactions. The technical core involves modeling the conditional intensity function $\lambda(e)$ of a heterogeneous event $e = (u, v, r, t)$, which characterizes the instantaneous arrival rate of a type- r relation between

source node u and target node v at time t :

$$\lambda(e) = \mu_r(u, v) + \sum_{e_h: t_h < t} \alpha(e_h, e) \cdot \kappa(t - t_h), \quad (3)$$

where $\mu_r(\cdot)$ represents the base rate, $\alpha(\cdot)$ measures the excitation influence from a historical event e_h , and $\kappa(\cdot)$ is a kernel function capturing the temporal decay influence. HPGE [Ji *et al.*, 2021] integrates a heterogeneous evolving attention mechanism to distinguish fine-grained excitation patterns between different-typed historical and current events, while utilizing temporal importance sampling to extract representative interactions efficiently. HDGAN [Li *et al.*, 2020] and THINE [Huang *et al.*, 2021] further leverage meta-paths to constrain candidate sets and employ hierarchical attention to jointly model semantic relevance and structural influence. SemE [Zhou *et al.*, 2023] abstracts chronological meta-path instances as semantic units driven by an attention-Hawkes process. Besides, TeSa [Wang *et al.*, 2025b] employs a trajectory-based neural TPP to encode node interaction sequences independently, enhancing scalability and long-term pattern capture.

Despite computational efficiency, embedding-based methods remain constrained by predefined proximity measures that limit representation capacity, shallow architectures that preclude higher-order structural modeling, and the absence of task-specific gradients that impedes downstream transfer.

3.2 GNN-based

GNN-based methods extend message-passing to DHGs by incorporating heterogeneous edge types and temporal evolution, predominantly operating at the snapshot granularity. Based on the mechanism for integrating type-aware propagation, we categorize these methods into: (1) relation-specific GNNs that decompose each snapshot into type-specific subgraphs with separate transformation matrices; (2) meta-structure-guided GNNs that constrain propagation along predefined meta-paths or meta-graphs to capture high-order semantic patterns; and (3) recurrent-based GNNs that couple spatial aggregation with temporal state evolution through RNN units.

Relation-specific GNNs

Relation-specific GNNs represent a dominant paradigm for modeling DHGs by adopting a divide-and-conquer strategy, decomposing complex heterogeneous interactions into multiple relation-specific subgraphs across different snapshots. These methods typically employ hierarchical message-passing and aggregation to disentangle the influence of different relation types while maintaining temporal consistency:

$$\mathbf{h}_v^{(t)} = \text{AGG}_{\text{time}} \left(\left\{ \text{AGG}_{\text{type}} \left(\{\mathbf{h}_u^{(t)}\}_{u \in \mathcal{N}_r^{(t)}(v)} \right) \right\}_{r \in \mathcal{R}} \right), \quad (4)$$

where $\mathcal{N}_r^{(t)}(v)$ denotes the neighbors of node v under relation r at time t . DyHAN [Yang *et al.*, 2020b] and HTGNN [Fan *et al.*, 2022] implement this via hierarchical attention, aggregating first across neighbors of the same type, then across types, and finally over time. Some approaches also incorporate hypergraph structures. For example, HTHGN [Liu *et al.*, 2025a] generalizes the receptive field to high-order group interactions through heterogeneous hyperedge construction and expansion to preserve local connectivity while capturing group dynamics.

Despite their expressiveness, such hierarchical architectures are difficult to scale, as separate message-passing over neighbor types and historical snapshots incurs substantial computational and memory overhead. Moreover, repeated attention-based aggregation across both type and time dimensions potentially leads to over-smoothing.

Meta-structure-guided GNNs

Meta-structure-guided GNNs predefine meta-paths or meta-graphs based on specific semantics and constrain message-passing only between nodes reachable from these structures, thus explicitly modeling semantic dependencies. Formally, given a snapshot $\mathcal{G}^{(t)}$ and a meta-structure $S \in \{\mathcal{P}, \mathcal{M}\}$, the representation of node v is computed by aggregating messages from its meta-structure-induced context $\mathcal{N}_s^{(t)}(v)$:

$$\mathbf{h}_v^{(t)} = \text{AGG}_{\text{time}} \left(\left\{ \text{AGG}_{\text{meta}} \left(\{\mathbf{h}_u^{(t)}\}_{u \in \mathcal{N}_s^{(t)}(v)} \right) \right\}_{s \in S} \right). \quad (5)$$

HDGAN [Li *et al.*, 2020] constrains propagation by enumerated meta-paths and treats each meta-path as an independent semantic channel, enabling additive composition of path-level semantics via attention-based fusion. DHANE [Li *et al.*, 2024b] merges multiple paths into a unified meta-graph structure and performs message-passing at the meta-graph level, thereby capturing cross-path contextual dependencies in a single propagation process. Furthermore, MG-DVD [Liu *et al.*, 2021] pre-enumerates frequent meta-graphs as semantic templates, under which node or graph representations are aggregated and then aggregated by template-level attention.

While semantically interpretable, these methods suffer from combinatorial explosion when enumerating meta-structures over rich type schemas, and their static semantic templates cannot adapt to temporal shifts in which meta-structures become salient at different evolutionary stages.

Recurrent-based GNNs

Recurrent-based GNNs capture the temporal evolution of node representations by maintaining historical states that are recur-

sively updated across snapshots or interactions. The core mechanism employs a heterogeneous GNN to encode structural information at each snapshot, coupled with recurrent units to evolve the node memory $\mathbf{M}^{(t)}$:

$$\mathbf{m}_v^{(t)} = f_{\text{RNN}} \left(\mathbf{m}_v^{(t-1)}, f_{\text{HGNN}} \left(\mathcal{G}^{(t)}, \mathbf{X}^{(t)} \right) \right), \quad (6)$$

where $\mathbf{m}_v^{(t)}$ denotes a recurrent memory state of node v at time t , and $f_{\text{HGNN}}(\cdot)$ encodes the heterogeneous structure and features of snapshot $\mathcal{G}^{(t)}$. DyHATR [Xue *et al.*, 2021] employs a hierarchical attention mechanism as the encoder to capture heterogeneity within each snapshot, followed by an RNN and temporal self-attention to model sequential dependencies. DURENDAL [Dileo *et al.*, 2023] generalizes this paradigm by investigating the structural placement of the recurrent update relative to the heterogeneous aggregation, proposing both Aggregate-Then-Update and Update-Then-Aggregate schemes to balance memory efficiency and relational temporal dynamics. Unlike the above methods that separate structural encoding and temporal updating, HGN2T [Liu *et al.*, 2024] proposes a plug-and-play coupling mechanism that tightly integrates a static heterogeneous GNN with a recurrent module, allowing the temporal evolution of node representations to be conditioned jointly on heterogeneous graph structure and relation-specific dynamics.

In summary, GNN-based methods enable end-to-end learning with expressive message-passing, yet remain constrained by hierarchical decomposition that prevents parallelization, snapshot discretization that discards fine-grained temporal signals, and computational complexity that scales multiplicatively with relation types and temporal depth.

3.3 Transformer-based

Transformer-based approaches represent the most recent paradigm shift in DHG representation learning, which leverage architectural advances that have driven breakthroughs across deep learning. By unifying DHG modeling via self-attention, they bypass the hierarchical decomposition inherent to GNN-based methods while enabling parallelized training. Three paradigms operate at distinct granularities: (1) structure-oriented methods, which attend over spatial neighborhoods at snapshot granularity with type-parameterized weights and relative time encodings; (2) interaction-oriented methods, which model each node’s event history at interaction granularity through sequence-level attention; and (3) LLM-enhanced methods, which inject external semantic knowledge from LLMs to complement structural representations.

Structure-oriented Transformer

Structure-oriented Transformers unify heterogeneous and temporal modeling within a single attention mechanism, departing from hierarchical GNNs that aggregate separately across relation types and time steps. The core paradigm computes attention over the full dynamic heterogeneous neighborhood, where attention weights jointly depend on source/target node types, edge type, and relative temporal position:

$$\mathbf{h}_v^{(l)} = \sum_{u \in \mathcal{N}(v)} \alpha_{vu}^{(r)} \cdot f_{\text{MSG}} \left(\mathbf{h}_u^{(l-1)}, f_{\text{RTE}}(\Delta t) \right), \quad (7)$$

Category	Method	Granularity	Core Technique	Task	Code
Embedding-based	Change2vec [Bian <i>et al.</i> , 2019]	Dis.	Meta-path RW with Incremental Update	Clust.	Link
	THINE [Huang <i>et al.</i> , 2021]	Con.	Chronological Meta-path Walk with Hawkes Process	NC/LP/TLP	Link
	DyHINE [Xie <i>et al.</i> , 2021b]	Dis.	Relation-specific RW with Att.	TLP	-
	HPGE [Ji <i>et al.</i> , 2021]	Con.	Relation-specific Att. with TPP	NC/TLP	Link
	M-DHIN [Fang <i>et al.</i> , 2022]	Dis.	Meta-graph Sampling with LSTM	NC/LP/TLP	-
	DyHNE [Wang <i>et al.</i> , 2022b]	Dis.	Meta-path Proximity with Eigen Perturbation	NC/RP/TLP	Link
	LIME [Peng <i>et al.</i> , 2022]	Dis.	Meta-path RW with Recursive NN	NC/Clust.	Link
	SemE [Zhou <i>et al.</i> , 2023]	Con.	Semantic Meta-path Walk with Hawkes Att.	NC/RP/TLP	Link
	H2TNE [Bai <i>et al.</i> , 2022]	Con.	Relation-specific Walk with Hyperbolic Embedding	NC/LP/TLP	Link
GNN-based	DHANE [Li <i>et al.</i> , 2024b]	Dis.	Meta-graph Att. with Online Update	NC	Link
	HDGAN [Li <i>et al.</i> , 2020]	Con.	Meta-path GNN with TPP	NC/Clust.	-
	DyHAN [Yang <i>et al.</i> , 2020b]	Dis.	Relation-specific GNN with Temporal Att.	TLP	-
	DyHATR [Xue <i>et al.</i> , 2021]	Dis.	Relation-specific GNN with Temporal Att.	TLP	Link
	HTGNN [Fan <i>et al.</i> , 2022]	Dis.	Relation-specific GNN with Temporal Modeling	NR/TLP	Link
	DURENDAL [Dileo <i>et al.</i> , 2023]	Dis.	Relation-specific GNN with RNN	LP/TLP	Link
	HGN2T [Liu <i>et al.</i> , 2024]	Dis.	Relation-specific GNN with GRNN	LP/TLP	Link
	HTHGN [Liu <i>et al.</i> , 2025a]	Dis.	Relation-specific GNN with Temporal Att.	LP/TLP	Link
	TeSa [Wang <i>et al.</i> , 2025b]	Con.	Relation-specific GNN with TPP	LP/TLP	-
Transformer-based	HGT [Hu <i>et al.</i> , 2020]	Con.	Relation-specific Transformer with RTE	TLP	Link
	HT-Trans [Ruan <i>et al.</i> , 2022]	Con.	Relation-specific Transformer Encoder	LP/TLP	-
	DHGAS [Zhang <i>et al.</i> , 2023b]	Dis.	Relation- and Time-aware Self-Att.	NC/NR/RP/LP/TLP	Link
	THGAT [Zhang <i>et al.</i> , 2023a]	Con.	Relation-specific Att. with RTE	NC/Clust./TLP	Link
	STHN [Li <i>et al.</i> , 2023a]	Con.	Structural-Temporal Transformer	LP/TLP	Link
	CasMLN [Wang <i>et al.</i> , 2024a]	Dis.	LLM-enhanced Relation-specific Transformer	NC/NR/RP/LP/TLP	Link
	HTGformer [Wang, 2025]	Dis.	LLM-enhanced Transformer for DHG	NC/NR/RP/LP/TLP	-
	MIGNN [Yue <i>et al.</i> , 2025]	Con.	Att.-based GNN with GRU	LP/TLP	-
Application-oriented	HTGT [Fan <i>et al.</i> , 2021]	Dis.	Relation-specific Att. with RTE	Cybersecurity	Link
	ATHITD [Qi <i>et al.</i> , 2025]	Dis.	Relation-specific Att. with RTE	Cybersecurity	-
	HetETA [Hong <i>et al.</i> , 2020]	ST.	Heterogeneous GNN with Causal Conv.	Traffic	Link
	STHGFormer [Li <i>et al.</i> , 2024a]	ST.	Relation-specific Att. with Temporal Att.	Traffic	-
	SUPA [Wu <i>et al.</i> , 2023]	Con.	Meta-path RW with Memory Update	Recomm.	Link
	DHGP [Li <i>et al.</i> , 2024c]	Dis.	Relation- and Time-aware Att.	Recomm.	Link
	DHGRNN [Wang <i>et al.</i> , 2025a]	ST.	Relation-specific Att. with GRU	Multimodal	-
	HDGR [Dai <i>et al.</i> , 2025]	Dis.	Relation-specific GConv. with Reconstruction	Multimodal	-
	SI-HDGN [Xu <i>et al.</i> , 2022]	Con.	Relation-specific RW with GRU	Diffusion	Link
	THGNN [Li <i>et al.</i> , 2023c]	Con.	Relation-specific Att. with LSTM	Diffusion	Link

Table 1: Comparison of DHG representation learning methods. Task abbreviations: NC (node classification), LP/TLP (link / temporal link prediction), RP (relation prediction), NR (node regression), Clust. (node clustering). Temporal granularity: Dis. (discrete-time), Con. (continuous-time), ST. (spatio-temporal). Technique abbreviations: RW (random walk), Att. (attention), GConv. (graph convolution), GRNN (graph recurrent neural network), TPP (temporal point process), RTE (relative time encoding).

where $\alpha_{vu}^{(r)}$ denotes the heterogeneous and temporal attention weight parameterized by the meta-relation triplet $(\phi(u), r, \phi(v))$ and relative temporal encoding $f_{\text{RTE}}(\Delta t)$. HGT [Hu *et al.*, 2020] pioneers this paradigm via triplet-based parameter decomposition and sinusoidal relative temporal encoding (RTE). HT-Trans [Ruan *et al.*, 2022] and DHGAS [Zhang *et al.*, 2023b] extend this with full Transformer encoders and neural architecture search for attention patterns. STHGFormer [Li *et al.*, 2024a] further refines this by integrating spatio-temporal positional encodings directly into the attention scores.

Despite unifying attention for DHG modeling, structure-oriented Transformers face scalability bottlenecks as attention over full neighborhoods across all relation types incurs quadratic complexity, and fine-grained type parameterization risks overfitting on infrequent meta-relations.

Interaction-oriented Transformer

Interaction-oriented Transformers diverge from neighbor-centric structural aggregation by modeling the continuous stream of edge events as a temporal sequence, where self-attention operates over a node’s chronological interaction history rather than its spatial neighborhood. The technical core

is to maintain a historical event sequence $\mathcal{H}_v = \{e_1, \dots, e_n\}$ for each node, where each event $e_i = (u_i, v, r_i, t_i)$ carries heterogeneous types and continuous timestamps, and to compute the node representations via sequence-level attention:

$$\mathbf{h}_v^{(t)} = \sum_{i=1}^{|\mathcal{H}_v|} \alpha_i^{(t)} \cdot (f_{\text{MSG}}(u_i, r_i) + f_{\text{RTE}}(\Delta t)), \quad (8)$$

where $\alpha_i^{(t)}$ denotes the attention weight at time t over the i -th historical event, $f_{\text{MSG}}(\cdot)$ encodes heterogeneous type information, and $f_{\text{RTE}}(\cdot)$ provides relative temporal position encoding. STHN [Li *et al.*, 2023a] introduces a unified link encoder that integrates type encoding and relative time encoding, and devises a patching technique that segments long event sequences into fixed-length patches to reduce complexity from quadratic to linear. MIGNN [Yue *et al.*, 2025] further employs a dual-window strategy to separately model short-term fine-grained dynamics and long-term distributional patterns through multi-scale temporal aggregation.

This paradigm excels at fine-grained temporal modeling by capturing event-level dynamics, enabling precise tracking of evolving interaction patterns. However, it incurs memory

overhead for maintaining per-node event histories and often underutilizes local structural context.

LLM-enhanced Transformer

LLM-enhanced Transformers integrate LLMs to inject external semantic knowledge into DHG learning, addressing the limitations of purely structural methods in capturing implicit type-level properties and domain-specific evolution patterns. Since different node types exhibit distinct temporal dynamics—for example, papers grow monotonically while e-commerce interactions show periodicity—LLMs provide complementary semantic priors that are difficult to learn from topology alone. The core paradigm fuses LLM-derived semantic embeddings with structural representations at multiple granularities:

$$\mathbf{h}_v^{(l)} = \sum_{u \in \mathcal{N}(v)} f_{\theta} \left(\mathbf{h}_u^{(t)}, f_{\text{LLM}}(p, \tau(\phi(v), \mathcal{G})) \right), \quad (9)$$

where p indicates the instruction prompt for the current graph context, and $\tau(\phi(v), \mathcal{G})$ denotes the textual description of node type and graph-level characteristics. The fusion function $f_{\theta}(\cdot)$ integrates these complementary views to produce semantically-enriched node embeddings. CasMLN [Wang *et al.*, 2024a] pioneers this paradigm by constructing structured prompts for both node types and the overall graph, then modulating aggregated representations with LLM-derived embeddings through element-wise product to compensate for skewed heterogeneous data distributions. HTGformer [Wang, 2025] aligns semantic and structural spaces through instruction tuning, enabling the model to interpret dynamic heterogeneous event patterns as textual instructions, shifting the paradigm from structural encoding towards semantic reasoning.

To summarize, Transformer-based methods unify DHG modeling within attention mechanisms, avoiding hierarchical decomposition. However, they remain constrained by quadratic complexity that limits scalability, additive fusion of type and temporal encodings rather than intrinsic coupling, the absence of pre-training objectives that precludes cross-domain transfer, and underexplored tokenization strategies for heterogeneous node/edge types and temporal granularities.

3.4 Application-oriented

Application-oriented approaches formulate domain-specific problems as DHG, serving as a unifying substrate where heterogeneous entities and relations evolve over time. Here we summarize representative applications in the following fields.

Cybersecurity

Cybersecurity addresses adversarial behavior detection where attackers deliberately fragment malicious actions across entity types and extended time horizons, rendering individual activities benign in isolation yet collectively revealing coordinated attacks. HTGT [Fan *et al.*, 2021] models app-market-developer relations through a heterogeneous temporal Transformer that iteratively aggregates spatial dependencies with historical sequences, jointly capturing malware propagation and evolution. MG-DVD [Liu *et al.*, 2021] translates API event streams into dynamic graphs with discriminative meta-graphs, enabling real-time variant detection via dynamic walks without full retraining. ATHITD [Qi *et al.*, 2025] introduces

temporal neighbors within sliding windows for short-term evolution, while Transformer modules learn long-term drift to highlight anomalous time periods. TimeSAGE [Shekhar *et al.*, 2020] incorporates time-decayed edge weights and temporal random walks to resolve fraudulent identities drifting over subscription lifecycles. These methods require accumulated behavioral evidence, enabling long-horizon attack detection but limiting responsiveness to zero-day exploits. While these methods effectively detect known attack patterns, adapting to adversaries’ evolving tactics remains an open challenge.

Traffic Forecasting

Traffic forecasting predicts dynamics across infrastructure where road segments follow smooth diurnal periodicity, intersections exhibit abrupt phase-dependent transitions, and transit hubs aggregate multimodal demand with distinct volatility patterns that uniform embeddings would collapse. HetETA [Hong *et al.*, 2020] constructs multi-relational networks with trajectory sequences, employing parallel temporal convolutions for recent, daily, and weekly patterns before graph propagation. STHGFormer [Li *et al.*, 2024a] explicitly distinguishes segments from turn nodes via heterogeneous spatial embeddings, with a unified Transformer capturing road-turn interdependencies. REGNN [Luo *et al.*, 2020] dynamically constructs event-centric graphs encoding ride requests with supply-demand context, enabling inductive generalization through transferable demand patterns. Although effective in capturing spatio-temporal patterns on observed road networks, generalization to unseen topologies or non-recurrent events remains limited.

Recommendation

Recommendation infers preferences from behavioral signals carrying heterogeneous semantic weight and temporal dynamics: clicks indicate transient interest with rapid decay, purchases reflect committed preferences with lasting influence, and reviews encode explicit sentiment requiring delayed integration. DHIMN [Xie *et al.*, 2021a] maintains type-specific memory states with two-level attention that weights items within behavior types then aggregates across types. DHGP [Li *et al.*, 2024c] introduces periodic prompts aligning cyclic interests with temporal context through prompt embeddings modulating basket representations. SUPA [Wu *et al.*, 2023] addresses neighborhood disturbance via sampling influenced subgraphs, updating endpoints, and propagating incrementally without full retraining. Despite effectively capturing type-specific preference, these methods remain limited under cold-start conditions with sparse heterogeneous interactions.

Multimodal Learning

Multimodal learning integrates data streams whose semantic alignments shift as context evolves, particularly when modalities exhibit asynchronous sampling rates or intermittent availability that preclude fixed correspondence assumptions. HSSHG [Wang *et al.*, 2024b] constructs spatio-temporal graphs where co-occurrence and consistency constrain adjacency, using plot summaries and locations as priors to weight edges for video question answering. DHGRNN [Wang *et al.*, 2025a] fuses spatial-temporal and spatial-spectral streams, with graph Transformers modeling channel heterogeneity and evolving convolutions adapting to missing EEG channels as

structure changes. HDGR [Dai *et al.*, 2025] builds dynamic intra-modal graphs with bipartite cross-modal graphs capturing semantic transitivity for 3D retrieval. While effective in fusing multimodal data with dynamic alignments, robustness to missing or noisy modalities remains underexplored.

Information Diffusion

Information diffusion examines propagation across heterogeneous actors governed by type-specific dynamics. For example, citations accrue monotonically, evidencing durable scholarly influence; social endorsements decay rapidly in accordance with attention cycles; and financial contagion transmits through correlation structures that shift abruptly under changing market regimes. HINTS [Jiang *et al.*, 2021] addresses cold-start prediction by imputing pseudo-historical trajectories via network alignment, converting embeddings into citation model parameters. SI-HDGNN [Xu *et al.*, 2022] models academic influence via temporal-attentive aggregation over directed author-paper-venue graphs. NetCycle+ [Xiong *et al.*, 2018] incorporates life-cycle stages, recognizing that nodes at different evolutionary phases exhibit distinct patterns. THGNN [Li *et al.*, 2023c] constructs daily correlation graphs with heterogeneous attention for financial contagion. Despite these efforts, unifying diffusion dynamics across actor types with disparate temporal scales remains open.

4 Open-source Datasets and Benchmarks

This section summarizes commonly used open-source DHG datasets, covering their application domains, numbers of node and edge types, number of time snapshots/timestamps, availability of node labels, and data source links. The datasets are ordered in descending usage frequency in Table 2.

Note that although many studies nominally rely on the same public sources, they frequently employ inconsistent graph construction rules, temporal discretizations, and evaluation protocols, which undermine reproducibility and cross-method comparability. To mitigate this, TGB 2.0 [Gastinger *et al.*, 2024] introduces standardized DHG datasets with unified pre-processing and evaluation pipelines, providing a more consistent foundation for benchmarking and facilitating future extensions. Nevertheless, the current release primarily targets link prediction and includes only a relative limited set of baseline methods. Prevailing fixed-split evaluation protocols that test solely on final-period edges fail to capture DHG

evolution and tend to overestimate. Extending task coverage, establishing incremental learning protocols, and developing comprehensive benchmarks remain critical for future research.

5 Conclusion and Future Directions

This survey presents the first systematic review of DHG representation learning. We provide a generalized definition of DHGs, establish an algorithm-centric taxonomy, and summarize key modeling biases. Our analysis of representative methods and benchmarks reveals important limitations in current modeling mechanisms and evaluation protocols. Despite recent advances toward foundation models [Liu *et al.*, 2025b], three critical challenges in DHG representation learning persist: (1) Prohibitive computational complexity hinders scaling to Web-scale graphs, precluding systematic investigation of scaling laws; (2) Current task-specific paradigms lack pre-training frameworks for learning transferable representations across heterogeneous schemas and temporal distributions; (3) Limited interpretability impedes high-stakes deployment and obscures whether models capture causal mechanisms versus spurious correlations. We conclude by identifying three progressive research frontiers to address these challenges that will catalyze future advances in this rapidly evolving field.

Efficiency. The quadratic complexity of current DHG learning methods [Hu *et al.*, 2020; Li *et al.*, 2024a] that explicitly couple message-passing with temporal aggregation fundamentally restricts their applicability to Web-scale DHGs. While recent efforts have attempted to alleviate this bottleneck through partial linearization strategies such as event patching [Li *et al.*, 2023a] or incremental propagation [Fang *et al.*, 2022], the expressiveness-efficiency trade-off persists. Thus, a critical open problem is to explore efficient alternatives, such as kernel-approximation-based variants of linear attention mechanisms or sparsification strategies for DHGs, as a prerequisite for investigating the scaling laws of DHG representation learning.

Generalizability. Current DHG methods are largely confined to task-specific supervised training, failing to exhibit transfer capabilities seen in vision and language foundation models. Although preliminary attempts have integrated LLM-derived semantics [Wang *et al.*, 2024a] or unified attention mechanisms across heterogeneous types [Wang, 2025], the field still lacks principled pre-training objectives for DHGs. Future research could develop pre-training paradigms such as self-supervised tasks, combined with transfer learning, few-shot adaptation, and continuous learning, to enable model generalizability across multi-domain evolutionary patterns.

Trustworthiness. As DHG models increasingly inform high-stakes decision-making, interpretability is imperative for ensuring trustworthiness. Current approaches [Li *et al.*, 2023b] emphasize post-hoc subgraph identification but lack faithfulness guarantees and causal grounding. Future work should prioritize inherently interpretable architectures, such as attention or bottleneck layers projecting onto human-interpretable concepts, and integrate structural causal models with DHG learning to distinguish genuine causal relationships from spurious correlations across heterogeneous entities and time.

	Domain	Dataset	$ A $	$ R $	$ T $	Label	URL
Datasets	Academic	DBLP	4	3	10	✓	Link
	Academic	AMiner	4	3	10	✓	Link
	Review	Yelp	4	3	10	✓	Link
	Social	Twitter	1	3	7	×	Link
	Epidemic	COVID	2	3	304	✓	Link
	Recomm.	MovieLens	3	3	25,864	×	Link
	Recomm.	Ecomm	2	4	11	×	Link
	Academic	MAG	4	4	10	✓	Link
	Recomm.	Alibaba	2	4	11	×	Link
	Academic	Math-Overflow	1	3	11	×	Link
TGB 2.0	Software	software	4	14	689,549	×	Link
	Social	forum	2	2	2,558,457	×	Link
	Software	github	4	14	2,510,415	×	Link
	Recomm.	myket	2	2	14,828,090	×	Link

Table 2: Overview of Open-source DHG Datasets and Benchmark.

Acknowledgments

Pengfei Jiao was partially supported by the National Natural Science Foundation of China under Grant No. 62372146, the Zhejiang Province Key R&D Program Project under Grants No. 2024C01212 and 2025C01023, and the Zhejiang Provincial Key Laboratory for Sensitive Data Security Protection and Confidentiality Management under Grant No. 2024E10048. Jie Yin was partially supported by the Australian Research Council under Grant No. DP250100871.

References

- [Bai *et al.*, 2022] Qijie Bai, Jiawen Guo, Haiwei Zhang, Changli Nie, Lin Zhang, and Xiaojie Yuan. H2tne: Temporal heterogeneous information network embedding in hyperbolic spaces. In *ISWC 2022*, pages 179–195, 2022.
- [Barros *et al.*, 2021] Claudio D. T. Barros, Matheus R. F. Mendonça, Alex B. Vieira, and Artur Ziviani. A survey on embedding dynamic graphs. *ACM Comput. Surv.*, 55(1), 2021.
- [Bian *et al.*, 2019] Ranran Bian, Yun Sing Koh, Gillian Dobbie, and Anna Divoli. Network embedding and change modeling in dynamic heterogeneous networks. In *SIGIR*, pages 861–864, 2019.
- [Dai *et al.*, 2025] Yue Dai, Yifan Feng, Nan Ma, Xibin Zhao, and Yue Gao. Cross-modal 3d shape retrieval via heterogeneous dynamic graph representation. *TPAMI*, 47(4):2370 – 2387, 2025.
- [Dileo *et al.*, 2023] Manuel Dileo, Matteo Zignani, and Sabrina Gaito. Durendal: Graph deep learning framework for temporal heterogeneous networks. In *NeurIPS*, 2023.
- [Fan *et al.*, 2021] Yujie Fan, Mingxuan Ju, Shifu Hou, Yanfang Ye, Wenqiang Wan, Kui Wang, Yinming Mei, and Qi Xiong. Heterogeneous temporal graph transformer: An intelligent system for evolving android malware detection. In *KDD*, pages 2831–2839, 2021.
- [Fan *et al.*, 2022] Yujie Fan, Mingxuan Ju, Chuxu Zhang, and Yanfang Ye. Heterogeneous temporal graph neural network. In *SDM*, pages 657–665, 2022.
- [Fang *et al.*, 2022] Yang Fang, Xiang Zhao, Peixin Huang, Weidong Xiao, and Maarten de Rijke. Scalable representation learning for dynamic heterogeneous information networks via metagraphs. *TOIS*, 40(4):64:1–64:27, 2022.
- [Fey *et al.*, 2024] Matthias Fey, Weihua Hu, Kexin Huang, Jan Eric Lenssen, Rishabh Ranjan, Joshua Robinson, Rex Ying, Jiaxuan You, and Jure Leskovec. Position: Relational deep learning - graph representation learning on relational databases. *ICML*, 235:13592–13607, 2024.
- [Gastinger *et al.*, 2024] Julia Gastinger, Shenyang Huang, Mikhail Galkin, Erfan Loghmani, Ali Parviz, Farimah Pourasaei, Jacob Danovitch, Emanuele Rossi, Ioannis Koutis, Heiner Stuckenschmidt, Reihaneh Rabbany, and Guillaume Rabusseau. Tgb 2.0: A benchmark for learning on temporal knowledge graphs and heterogeneous graphs. In *NeurIPS*, volume 37, pages 140199–140229, 2024.
- [Hong *et al.*, 2020] Huiting Hong, Yucheng Lin, Xiaoqing Yang, Zang Li, Kung Fu, Zheng Wang, Xiaohu Qie, and Jieping Ye. Heterogeneous information network embedding for estimating time of arrival. In *KDD*, 2020.
- [Hu *et al.*, 2020] Ziniu Hu, Yuxiao Dong, Kuansan Wang, and Yizhou Sun. Heterogeneous graph transformer. In *WWW*, pages 2704–2710, 2020.
- [Huang *et al.*, 2021] Hong Huang, Ruize Shi, Wei Zhou, Xiao Wang, Hai Jin, and Xiaoming Fu. Temporal heterogeneous information network embedding. In *IJCAI*, pages 1470–1476, 2021.
- [Ji *et al.*, 2021] Yugang Ji, Tianrui Jia, Yuan Fang, and Chuan Shi. Dynamic heterogeneous graph embedding via heterogeneous hawkes process. In *PKDD*, 2021.
- [Jiang *et al.*, 2021] Song Jiang, Bernard Koch, and Yizhou Sun. Hints: Citation time series prediction for new publications via dynamic heterogeneous information network embedding. In *WWW*, pages 3158–3167, 2021.
- [Jin *et al.*, 2024] Bowen Jin, Gang Liu, Chi Han, Meng Jiang, Heng Ji, and Jiawei Han. Large language models on graphs: A comprehensive survey. *TKDE*, 36(12):8622–8642, 2024.
- [Kazemi *et al.*, 2020] Seyed Mehran Kazemi, Rishabh Goel, Kshitij Jain, Ivan Kobyzev, Akshay Sethi, Peter Forsyth, and Pascal Poupard. Representation learning for dynamic graphs: A survey. *JMLR*, 21(70):1–73, 2020.
- [Li *et al.*, 2020] Qiuyan Li, Yanlei Shang, Xiuquan Qiao, and Wei Dai. Heterogeneous dynamic graph attention network. In *ICKG*, pages 404–411, 2020.
- [Li *et al.*, 2023a] Ce Li, Rongpei Hong, Xovee Xu, Goce Trajcevski, and Fan Zhou. Simplifying temporal heterogeneous network for continuous-time link prediction. In *CIKM*, pages 1288–1297, 2023.
- [Li *et al.*, 2023b] Jiazheng Li, Chunhui Zhang, and Chuxu Zhang. Heterogeneous temporal graph neural network explainer. In *CIKM*, pages 1298–1307, 2023.
- [Li *et al.*, 2023c] Yilin Li, Jiaqi Zhu, Congcong Zhang, Yi Yang, Jiawen Zhang, Ying Qiao, and Hongan Wang. Thgnn: An embedding-based model for anomaly detection in dynamic heterogeneous social networks. In *CIKM*, 2023.
- [Li *et al.*, 2024a] Guangyue Li, Zilong Zhao, Xiaogang Guo, Luliang Tang, Huazu Zhang, and Jinghan Wang. Towards integrated and fine-grained traffic forecasting: A spatio-temporal heterogeneous graph transformer approach. *Inform. Fusion*, 102, 2024.
- [Li *et al.*, 2024b] Hongbo Li, Wenli Zheng, Feilong Tang, Yitong Song, Bin Yao, and Yanmin Zhu. Dynamic heterogeneous attributed network embedding. *INS*, 662, 2024.
- [Li *et al.*, 2024c] Ru-Bin Li, Man-Sheng Chen, Xin-Yu Ding, Chang-Dong Wang, Sihong Xie, Shuangyin Liu, Min Chen, and Mohsen Guizani. Periodic prompt on dynamic heterogeneous graph for next basket recommendation. In *ICDM*, pages 747–752, 2024.
- [Liu *et al.*, 2021] Chen Liu, Bo Li, Jun Zhao, Ming Su, and Xu-Dong Liu. Mg-dvd: A real-time framework for malware

- variant detection based on dynamic heterogeneous graph learning. In *IJCAI*, pages 1512–1519, 2021.
- [Liu *et al.*, 2024] Huan Liu, Pengfei Jiao, Xuan Guo, Huaming Wu, Mengzhou Gao, and Jilin Zhang. Hgn2t: A simple but plug-and-play framework extending hgns on heterogeneous temporal graphs. *TBD*, 10(5):620–632, 2024.
- [Liu *et al.*, 2025a] Huan Liu, Pengfei Jiao, Mengzhou Gao, Chaochao Chen, and Di Jin. Heterogeneous temporal hypergraph neural network. In *IJCAI*, pages 3117–3125, 2025.
- [Liu *et al.*, 2025b] Jiawei Liu, Cheng Yang, Zhiyuan Lu, Junze Chen, Yibo Li, Mengmei Zhang, Ting Bai, Yuan Fang, Lichao Sun, Philip S. Yu, and Chuan Shi. Graph foundation models: Concepts, opportunities and challenges. *TPAMI*, 47:5023–5044, 2025.
- [Luo *et al.*, 2020] Wenjuan Luo, Han Zhang, Xiaodi Yang, Lin Bo, Xiaoqing Yang, Zang Li, Xiaohu Qie, and Jieping Ye. Dynamic heterogeneous graph neural network for real-time event prediction. In *KDD*, pages 3213–3223, 2020.
- [Peng *et al.*, 2022] Hao Peng, Renyu Yang, Zheng Wang, Jianxin Li, Lifang He, Philip S. Yu, Albert Y. Zomaya, and Rajiv Ranjan. Lime: Low-cost and incremental learning for dynamic heterogeneous information networks. *TC*, 71(3):628–642, 2022.
- [Qi *et al.*, 2025] Yinhao Qi, Chuyi Yan, Zehui Wang, Chen Zhang, Song Liu, Zhigang Lu, and Bo Jiang. Attention-based temporal heterogeneous graph neural network for insider threat detection. *COMPUT SECUR*, 157, 2025.
- [Ruan *et al.*, 2022] Beibei Ruan, Cui Zhu, and Wenjun Zhu. A link prediction model of dynamic heterogeneous network based on transformer. In *IJCNN*, pages 01–07, 2022.
- [Shekhar *et al.*, 2020] Shubhranshu Shekhar, Deepak Pai, and Sriram Ravindran. Entity resolution in dynamic heterogeneous networks. In *WWW*, pages 662–668, 2020.
- [Wang *et al.*, 2022a] Xiao Wang, Deyu Bo, Chuan Shi, Shao-hua Fan, Yanfang Ye, and Philip S Yu. A survey on heterogeneous graph embedding: methods, techniques, applications and sources. *TBD*, 9(2):415–436, 2022.
- [Wang *et al.*, 2022b] Xiao Wang, Yuanfu Lu, Chuan Shi, Ruijia Wang, Peng Cui, and Shuai Mou. Dynamic heterogeneous information network embedding with meta-path based proximity. *TKDE*, 34(3):1117–1132, 2022.
- [Wang *et al.*, 2024a] Fengyi Wang, Guanghui Zhu, Chunfeng Yuan, and Yihua Huang. Llm-enhanced cascaded multi-level learning on temporal heterogeneous graphs. In *SIGIR*, pages 512–521, 2024.
- [Wang *et al.*, 2024b] Ruomei Wang, Yuanmao Luo, Fuwei Zhang, Mingyang Liu, and Xiaonan Luo. Hsshg: Heuristic semantics-constrained spatio-temporal heterogeneous graph for videoqa. *TMM*, 26:11176 – 11190, 2024.
- [Wang *et al.*, 2025a] Jing Wang, Zhiyang Feng, Xiaojun Ning, Youfang Lin, Badong Chen, and Ziyu Jia. Two-stream dynamic heterogeneous graph recurrent neural network for multi-label multi-modal emotion recognition. *TAC*, 16(3):2396 – 2409, 2025.
- [Wang *et al.*, 2025b] Xin Wang, Jiawei Jiang, Xiao Yan, and Qiang Huang. Tesa: A trajectory and semantic-aware dynamic heterogeneous graph neural network. In *WWW*, pages 1305–1315, 2025.
- [Wang, 2025] Yili Wang. Htgformer: Heterogeneous temporal graph transformer. In *SIGIR*, pages 2550–2554, 2025.
- [Wu *et al.*, 2023] Cheng Wu, Chaokun Wang, Jingcao Xu, Ziwei Fang, Tiankai Gu, Changping Wang, Yang Song, Kai Zheng, Xiaowei Wang, and Guorui Zhou. Instant representation learning for recommendation over large dynamic graphs. In *ICDE*, pages 82–95, 2023.
- [Xie *et al.*, 2021a] Tao Xie, Yangjun Xu, Liang Chen, Yang Liu, and Zibin Zheng. Sequential recommendation on dynamic heterogeneous information network. In *ICDE*, pages 2105–2110, 2021.
- [Xie *et al.*, 2021b] Yuanzhen Xie, Zijing Ou, Liang Chen, Yang Liu, Kun Xu, Carl Yang, and Zibin Zheng. Learning and updating node embedding on dynamic heterogeneous information network. In *WSDM*, pages 184–192, 2021.
- [Xiong *et al.*, 2018] Yun Xiong, Yizhou Zhang, Xiangnan Kong, and Yangyong Zhu. Netcycle+: A framework for collective evolution inference in dynamic heterogeneous networks. *TKDE*, 30(8):1547–1560, 2018.
- [Xu *et al.*, 2022] Xovee Xu, Ting Zhong, Ce Li, Goce Trajceviski, and Fan Zhou. Heterogeneous dynamical academic network for learning scientific impact propagation. *KBS*, 238:107839, 2022.
- [Xue *et al.*, 2021] Hansheng Xue, Luwei Yang, Wen Jiang, Yi Wei, Yi Hu, and Yu Lin. Modeling dynamic heterogeneous network for link prediction using hierarchical attention with temporal rnn. *PKDD*, pages 282–298, 2021.
- [Yang *et al.*, 2020a] Carl Yang, Yuxin Xiao, Yu Zhang, Yizhou Sun, and Jiawei Han. Heterogeneous network representation learning: A unified framework with survey and benchmark. *TKDE*, 34(10):4854–4873, 2020.
- [Yang *et al.*, 2020b] Luwei Yang, Zhibo Xiao, Wen Jiang, Yi Wei, Yi Hu, and Hao Wang. Dynamic heterogeneous graph embedding using hierarchical attentions. In *ECIR*, pages 425 – 432, 2020.
- [Yue *et al.*, 2025] Chongjian Yue, Qiao Mi, and Lun Du. A unified framework for link prediction on heterogeneous temporal graph. In *APWeb 2024*, pages 310–322, 2025.
- [Zhang *et al.*, 2023a] Lin Zhang, Jiawen Guo, Qijie Bai, and Chunyao Song. Dynamic heterogeneous graph representation learning with neighborhood type modeling. *Neurocomputing*, 533(C):46–60, 2023.
- [Zhang *et al.*, 2023b] Zeyang Zhang, Ziwei Zhang, Xin Wang, Yijian Qin, Zhou Qin, and Wenwu Zhu. Dynamic heterogeneous graph attention neural architecture search. *AAAI*, 37(9):11307–11315, 2023.
- [Zhou *et al.*, 2023] Wei Zhou, Hong Huang, Ruize Shi, Xiran Song, Xue Lin, Xiao Wang, and Hai Jin. Temporal heterogeneous information network embedding via semantic evolution. *TKDE*, 35(12):13031–13042, 2023.