Scalable Semi-supervised Community Search via Graph Transformer on Attributed Heterogeneous Information Networks

Linlin Ding[1], **Zhaosong Zhao**[1], **Mo Li**[1*], **Yishan Pan**[1], **Xin Wang**[2], **Renata Borovica-Gajic**[3]

[1]Liaoning University, Shenyang, China
[2]Tianjin University, Tianjin, China
[3]The University of Melbourne, Melbourne, Australia

Abstract

Attributed heterogeneous information networks (AHINs) encode rich semantics through diverse node and edge types. Recent learning-based community search methods on AHINs have shown promising performance but face two major limitations: (i) difficulty scaling to large graphs due to memory-intensive neighbor-based propagation (e.g., GNNs and node-level attention), and (ii) reliance on explicit community-level labels, which are often unavailable or costly to obtain. To address these issues, we propose a scalable Semi-supervised Community Search framework on AHINs (SCSAH), enabling scalability and efficiency, while eliminating the need for community-level labels by leveraging readily available node classification labels. Specifically, we devise MvSF2Token to extract Multi-view Semantic Features (MvSFs) as compact subgraph-level tokens before training, significantly reducing model propagation complexity. We then design a View-Aware Semantic Graph Transformer (VASGhormer) to effectively encode MvSFs by capturing cross-view dependencies and fusing semantic features. The combination of MvSF2Token and VASGhormer ensures scalability, efficiency, and robust performance. Furthermore, we propose a novel View-Aware Contrastive Learner to train VASGhormer without requiring community-level supervision. Extensive experiments on five real-world datasets show that SCSAH significantly outperforms state-of-the-art methods, achieving 18.06% higher performance and 10.43× faster training. Our code is included in the supplementary material.

Introduction

Heterogeneous Information Networks (HINs) offer a powerful framework for modeling complex, multi-typed data and capturing rich semantic relationships across various domains, such as social networks, academic collaborations, and e-commerce platforms (Chen et al. 2022; Shi et al. 2017). Attributed HINs (AHINs) further extend this paradigm by integrating node attributes (Li et al. 2022b), thereby better reflecting the characteristics of real-world data. Within an AHIN, shown in Figure 1(a), meta-paths (Figure 1(b)) connect the same type nodes via predefined type sequences. These meta-paths not only express explicit semantics, e.g., the Author-Paper-Author (APA) path for co-authorship, but enable the transformation of an AHIN into a homogeneous graph (Figure 1(c)) for downstream analysis.

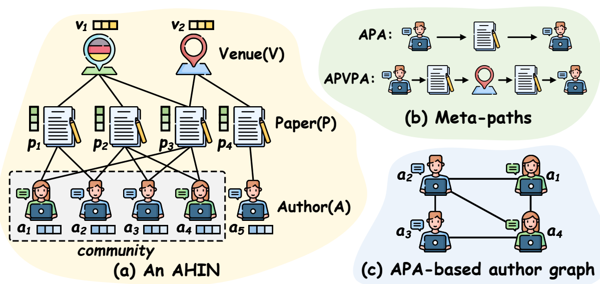

Figure 1: A toy AHIN from the DBLP network. (a) An AHIN composed of three types of nodes with their attributes, and the dashed boxes indicate potential community. (b) Examples of meta-path. (c) A homogeneous graph constructed from the AHIN based on the APA meta-path.

Community search (CS) is a fundamental problem in graph analysis that aims to identify a query-dependent, densely connected subgraph (Gou et al. 2023; Ye, Zhu, and Chen 2023). In recent years, the extension of CS to AHINs has garnered increasing attention (Wang et al. 2024b; Chen et al. 2024), driven by applications such as fraud detection (Zhong et al. 2020) and gene function prediction (Li et al. 2022a). Most existing CS methods on AHINs are algorithmic (Zhou et al. 2023; Liu et al. 2024; Wang et al. 2024b,d), relying on predefined structural constraints, such as the $(k, \mathcal{P})$-core, where k and $\mathcal{P}$ denote the degree and meta-path constraints, respectively. However, these algorithmic approaches typically overlook node attribute information, limiting their flexiblity and expressive power.

Recent studies have begun exploring machine learning (ML)-based approaches to achieve more flexible CS and enhanced feature modeling. A straightforward strategy is to transform an AHIN into a homogeneous graph via meta-paths and then apply existing homogeneous graph-based methods for CS. However, this transformation typically results in a loss of critical attribute and structural information inherent in AHINs. To address this limitation, recent approaches propose directly modeling AHINs by leveraging graph neural networks (GNNs) and attention mech-

*Corresponding Author: limo@lnu.edu.cn

anisms. For instance, ST-GNN (Li et al. 2024) is the first to utilize GNNs with self-supervised training to capture heterogeneous node features. FCS-HGNN (Chen et al. 2024) further integrates GNNs and attention mechanisms for multi-type CS under label supervision. Furthermore, CS-DAHIN (Song et al. 2024) focuses on dynamic AHINs, employing attention-based attribute scoring to assist CS.

Nevertheless, these advanced methods still face two major limitations: **i) Limited Scalability on Large-Scale Graphs**. All these methods rely on structure-based propagation (e.g., GNNs in ST-GNN and FCS-HGNN, or node-level attention in CS-DAHIN), which aggregate information from neighboring nodes and requires processing a large number of node embeddings. For large graphs, this is inefficient and results in substantial GPU memory overhead during training, as intermediate activations must be stored for backpropagation, fundamentally limiting their scalability. **ii) Reliance on Community Labels for Effective Training**. For example, CS-DAHIN simply treats nodes with same-classification as positives, which fails to capture community-level information (here, classification refers to the category of specific-type nodes, such as author nodes labeled as data mining, ML, etc.). ST-GNN relies on pseudo-labels for training, which may not accurately reflect actual community patterns, while FCS-HGNN depends on community labels that are often unavailable in practice. Thus, effective training strategies without relying on community labels remain underexplored.

To address these limitations, we propose a novel **S**emi-supervised **C**ommunity **S**earch framework over **AH**INs (SCSAH), which enables scalable and efficient CS on large graphs without relying on community labels. The SCSAH consists of three stages: 1) MvSF2Token, for the pre-construction of Multi-view Semantic Features (MvSFs), 2) offline training of the View-Aware Semantic Graph Transformer (VASGhormer), and 3) online community search.

Specifically, **for Limitation i**, we propose MvSF2Token, which extracts MvSFs based on meta-paths and graph structures, transforming them into subgraph-level tokens before training. As these tokens are much fewer than node-level tokens, they replace node-level tokens in model propagation, substantially reducing token volume and GPU memory overhead on large graphs during training. To further improve effectiveness and efficiency, we design VASGhormer, which incorporates a Zoom-Aware Transformer to capture cross-view dependencies and fuse semantic features. Together, MvSF2Token and VASGhormer enable our framework to achieve scalability, efficiency, and strong performance. **To address Limitation ii**, we design a View-Aware Contrastive Learner that achieves semi-supervised training by leveraging readily available node-level labels (e.g., the categories of author-type nodes) rather than costly community-level labels. Specifically, the learner combines semantic and unified contrastive losses, guided by a sampling strategy based on node categories, and incorporates node classification as an auxiliary task. Consequently, our approach achieves strong representation learning and model robustness while eliminating the need for community-level labels. Our main contributions are summarized as follows:

- We propose MvSF2Token to encode multi-view semantic features into subgraph-level tokens for downstream model use, effectively mitigating scalability issues.
- Building on this representation, we design VASGhormer to model cross-view and semantic dependencies among these tokens, enabling scalable and efficient propagation and training on large graphs.
- We introduce a View-Aware Contrastive Learner that enables semi-supervised training with the aid of node-level labels instead of community-level supervision.
- Experiments on five public datasets demonstrate that SCSAH outperforms state-of-the-art methods, achieving 18.06% higher performance and 10.43× faster training.

Related Work

Algorithmic Community Search on (A)HINs. Early CS studies on HINs focus on structure, often ignoring node attributes (Wang et al. 2020; Fang et al. 2020; Jiang et al. 2022a; Yang et al. 2022). For instance, the $(k, \mathcal{P})$-core model captures cohesiveness via meta-paths, facilitating cohesive community discovery (Wang et al. 2020). Keyword-centric methods further incorporate textual information into the search process (Qiao et al. 2021). Zhou et al. (Zhou et al. 2023) consider influence and multi-type nodes, introducing Pareto-optimality to improve search quality. With the growing adoption of AHINs, recent approaches introduce attribute-sensitive methods that combine structural and attribute cohesiveness for more effective search (Wang et al. 2024b,d). However, these methods still struggle to effectively model complex heterogeneous features.

ML-based Community Search. Recent years have seen an increasing adoption of ML methods for CS, owning to the powerful feature capture capabilities of neural networks. **For homogeneous graphs**, ICS-GNN (Gao et al. 2021) firstly formulates CS as a learning task using GNN, while COCLEP (Li et al. 2023) pioneers the integration of contrastive learning. Subsequent studies focus on scalable and feature-aware training (Wang et al. 2024c), as well as addressing label sparsity (Wang et al. 2024a). SLRL (Ni et al. 2025) further employs reinforcement learning to guide community generation. In contrast, there is limited research **on heterogeneous graphs**. ST-GNN (Li et al. 2024) is the first to apply GNNs with self-supervised training. FCS-HGNN (Chen et al. 2024) combines GNNs and attention to perform multi-type CS, while CS-DAHIN (Song et al. 2024) utilizes multi-level attention for attribute scoring over dynamic AHINs. However, these models are difficult to scale to large graphs, due to the high computational costs of graph-based propagation, and perform poorly when community labels are missing. Furthermore, due to inherent structural differences, models designed for homogeneous graphs cannot be directly applied to HINs, underscoring the need for further research on CS in heterogeneous settings.

Problem Definition

Definition 1 (Attributed Heterogeneous Information Network, AHIN)**.** An AHIN is a graph defined as $\mathcal{G} = (V, E, X, \varphi, \psi)$, where each node $v \in V$ and each edge

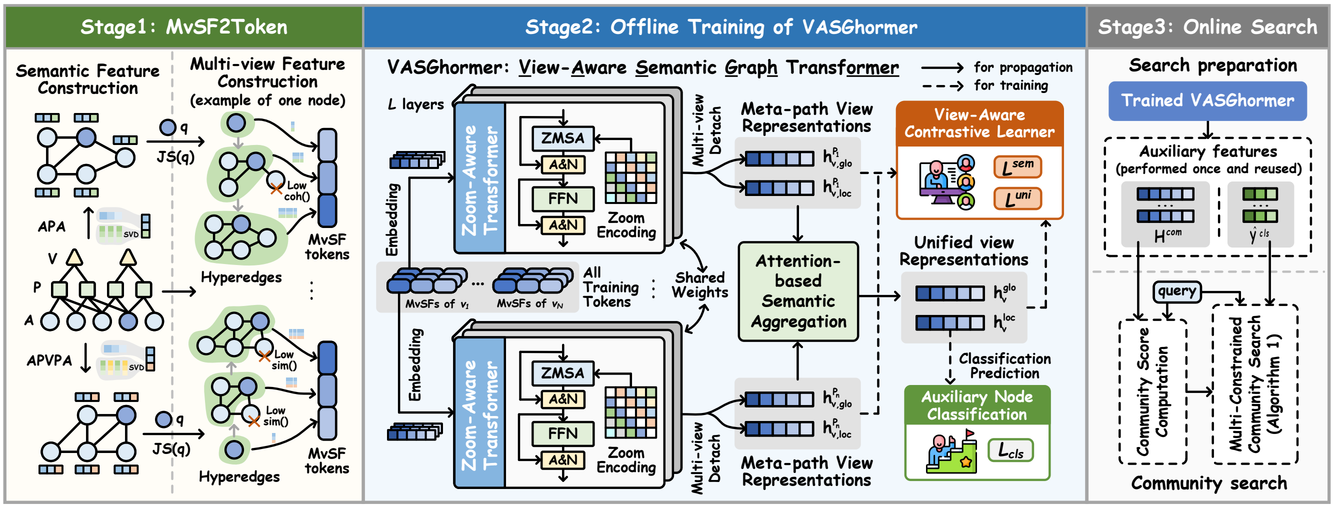

Figure 2: The overall framework of **SCSAH**. (1) **MvSF2Token**: extracts multi-view semantic features (MvSFs) from the AHIN, constructing compact subgraph-level tokens before training; (2) **Offline Training of VASGhormer**: the VASGhormer efficiently learns representations with a semi-supervised contrastive strategy; (3) **Online Search**: generates auxiliary features via the trained VASGhormer and performs Multi-Constrained Community Search.

$e \in E$ are associated with their specific type mapping functions $\varphi(v) : V \rightarrow \mathcal{A}$ and $\psi(e) : E \rightarrow \mathcal{R}$. Here, $\mathcal{A}$ and $\mathcal{R}$ refer to the sets of node and edge types, where $|\mathcal{A}|+|\mathcal{R}| > 2$. When $|\mathcal{A}| = |\mathcal{R}| = 1$, the graph degenerates into a homogeneous graph. $X^{(\mathcal{A}_i)} \subseteq X$ is the attribute set of $\mathcal{A}_i$ type nodes. $x_j^{(\mathcal{A}_i)} \in \mathbb{R}^{d_i}$ is the attribute of $v_j^{(\mathcal{A}_i)} \in V^{(\mathcal{A}_i)}$, where d_i is the dimension of the attributes of $\mathcal{A}_i$ type nodes.

Definition 2 (Meta-Path)**.** A meta-path $\mathcal{P}$ is defined as a path in the form of $\mathcal{A}_1 \xrightarrow{\mathcal{R}_1} \mathcal{A}_2 \xrightarrow{\mathcal{R}_2} \dots \xrightarrow{\mathcal{R}_l} \mathcal{A}_{l+1}$ which describes a composite relation between types $\mathcal{A}_1$ and $\mathcal{A}_{l+1}$.

Problem statement (Community Search on AHINs). Given an AHIN $\mathcal{G}$, a target type $\mathcal{A}_t$, a $\mathcal{A}_t$ type query node v_q, and a meta-path set $\mathcal{P}_S = \{\mathcal{P}_1, \dots, \mathcal{P}_M\}$, the task of Community Search on AHINs aims to identify a set of nodes C_q of the target type $\mathcal{A}_t$, where the nodes in C_q are densely intra-connected via meta-paths $\mathcal{P}_S$ and exhibit similar attributes.

Methodology

This section details the SCSAH framework, as shown in Figure 2. It comprises three stages: 1) MvSF2Token; 2) Offline Training of VASGhormer and 3) Online Search.

MvSF2Token: Feature Pre-construction

In CS tasks, efficiently extracting structural features is critical for overcoming the scalability limitations of GNNs and neighbor-dependent Transformers on large graphs. While prior work (Chen et al. 2023) achieves this by precomputing multi-hop neighbor representations for homogeneous graphs, structural features in AHINs are inherently more complex due to the presence of multi-type nodes and meta-paths. Consequently, existing pre-computation strategies are insufficient for AHINs. To address this, we propose **MvSF2Token**, a non-ML method for transforming Multi-view Semantic Features (MvSFs) from AHIN's structure into subgraph-level tokens before training, significantly reduces memory overhead of subsequent model training. It comprises two main stages: **Semantic Feature Construction** and **Multi-view Feature Construction**.

Semantic Feature Construction. A common way to unify feature dimensions across node types for semantic feature aggregation is padding, which, however, may dilute the contribution of target-type features. To mitigate this, we introduce a semantic feature extraction method based on Singular Value Decomposition (SVD).

Specifically, we first apply SVD to reduce the feature dimension of non-target-type nodes to a predefined dimension d_{oth}, which is smaller than that of target-type nodes. This ensures that, after concatenating features, target-node features retain their dominant influence. Given a meta-path $\mathcal{P}$, we then construct new semantic features for each target node. For a target node v, its semantic feature is computed as $x_v^{\mathcal{P}} = [x_v \,\|\, \frac{1}{|\mathcal{N}_{\mathcal{P}}^{\mathcal{G}}(v)|} \sum_{u \in \mathcal{N}_{\mathcal{P}}^{\mathcal{G}}(v)} \tilde{x}_u] \in \mathbb{R}^{d_0}$, where x_v is the original feature, $\tilde{x}_u$ denotes the SVD-reduced feature of a non-target node u, d_0 is the new concatenated dimension, and $\mathcal{N}_{\mathcal{P}}^{\mathcal{G}}(v)$ denotes the non-target-type neighbors of v along $\mathcal{P}$. For example, under the meta-path APVPA, $\mathcal{N}_{\mathcal{P}}^{\mathcal{G}}(v)$ includes P-type nodes connected to v, and V-type nodes connected to those P-type nodes. After processing all target-type nodes, we obtain the semantic features $X^{\mathcal{P}} \in \mathbb{R}^{N \times d_0}$, where N is the number of target-type nodes. Finally, by repeating the above process for different meta-paths, we construct a set of semantic feature matrices $\{X^{\mathcal{P}_1}, \dots, X^{\mathcal{P}_M}\}$ and their corresponding meta-path-based homogeneous graphs $\{\mathcal{G}^{\mathcal{P}_1}, \dots, \mathcal{G}^{\mathcal{P}_M}\}$ for subsequent stages.

Multi-view Feature Construction. Extracting meaningful community features is critical for CS tasks. Although simple aggregation of multi-hop neighbors (Chen et al. 2023; Wang et al. 2024a) enables efficient structural feature extraction, it may also incorporate weakly connected nodes, thereby undermining key community properties such as cohesion and feature similarity.

To address this, we propose a novel Multi-view Feature Construction method that selects nodes for each hop view based on a joint score, which enforces both structural and feature cohesion within the selected set for aggregation.

Specifically, given a meta-path $\mathcal{P}$, target node v, and hop limit K, for each hop $k \in \{1, \dots, K\}$, we first construct a k-hop neighbor subgraph $\mathcal{N}_k^{\mathcal{P}}(v)$ from $\mathcal{G}^{\mathcal{P}}$. For each node $u \in \mathcal{N}_k^{\mathcal{P}}(v)$, we compute a joint score: $\mathrm{JS}_k(u) = w_{\mathrm{s}}(u)\,\mathrm{sim}(v,u) + w_{\mathrm{c}}(u)\,\mathrm{coh}(u)$, where $\mathrm{sim}(v,u)$ is the cosine similarity between node features, and $\mathrm{coh}(u)$ is the Local Clustering Coefficient (Watts and Strogatz 1998) of u, measuring the structural cohesion. The weights $w_{\mathrm{s}}(u)$ and $w_{\mathrm{c}}(u)$ are defined as:

$$w_{\mathrm{s}}(u) = \frac{e^{\mathrm{coh}(v,u)}}{e^{\mathrm{sim}(v,u)} + e^{\mathrm{coh}(u)}}, \quad w_{\mathrm{c}}(u) = 1 - w_{\mathrm{s}}(u). \quad (1)$$

This equation adaptively balances structural cohesion and feature similarity, ensuring a high $\mathrm{JS}_k(u)$ only when both are strong (see Appendix for proof). Nodes with $\mathrm{JS}_k(u) \geq \theta$, where θ is cohesion threshold, are retained to form the k-hop view subgraph. Repeating this process up to K hops, yields K subgraphs form distinct views. Next, we encode v and its multi-view subgraph as hyperedges, forming a hypergraph $\mathcal{H}_v^{\mathcal{P}} \in \mathbb{R}^{N \times (K+1)}$, where each column represents one view. Finally, based on the semantic feature matrix $X^{\mathcal{P}}$, we compute the multi-view semantic features (MvSFs) for node v as $\mathbf{X}_v^{\mathcal{P}} \leftarrow (\mathcal{H}_v^{\mathcal{P}})^{\top} X^{\mathcal{P}} \in \mathbb{R}^{(K+1)\times d_0}$.

Notably, **MvSFs of all training nodes are constructed before the training phase**, reducing the number of tokens processed by the subsequent model from $\mathcal{O}(N)$ to $\mathcal{O}(K)$ while preserving rich structural features. The pseudocode for MvSF2Token is provided in the Appendix.

Design of VASGhormer

To effectively model interactions among MvSFs, a straightforward approach is to apply a standard Transformer architecture. However, conventional Transformers treat all input tokens equally and fail to capture cross-view relationships and fuse information from meta-paths, limiting their effectiveness for community-oriented tasks on AHINs. To address this, we propose the View-Aware Semantic Graph Transformer (VASGhormer), a tailored model designed to work collaboratively with MvSFs. VASGhormer enables scalable and efficient training while capturing cross-view community dependencies, making it particularly suitable for CS on AHINs. The model design is detailed as follows:

Token Embedding. Under each meta-path $\mathcal{P}$, we first project the feature tokens $\mathbf{X}_v^{\mathcal{P}}$ into a dense representation using a linear transformation, i.e., $\mathbf{H}_{\mathcal{P},v}^{(0)} = \mathbf{X}_v^{\mathcal{P}}\mathbf{W}_e$, where $\mathbf{W}_e \in \mathbb{R}^{d_0 \times d}$ is a learnable weight matrix.

Zoom-Aware Transformer. Transformer-based methods for graphs (Ying et al. 2021; Rampásek et al. 2022; Bar-Shalom, Bevilacqua, and Maron 2024) aim to capture latent relationships among tokens using structural priors like shortest path or edge weights. However, existing variants focus solely on node-level input tokens and fail to explicitly model interactions between tokens from different hops. Such interactions are crucial for integrating local and global information, which is essential to generate coherent representations of multi-view community features.

To model these cross-hop relations, we design the Zoom-Aware Transformer, which enhances Multi-Head Self-Attention via a Zoom-aware Multi-head Self-Attention (ZMSA). ZMSA captures progressive correlations among tokens across hop-level views by introducing a zoom-aware bias. Given two tokens $\mathbf{h}_{v,i}$ and $\mathbf{h}_{v,j}$ from the i-hop and j-hop views respectively, we define a zoom feature $\phi(i,j) = j - i$. A learnable scalar bias $\mathbf{b}_{\phi(i,j)}$ is associated with each zoom level, indicating a zoom-out when $\phi(i,j) > 0$ and a zoom-in when $\phi(i,j) < 0$. The (i,j)-element of the Query-Key product matrix in ZMSA is computed as:

$$\mathbf{A}^{zoom}(i,j) = \frac{(\mathbf{h}_{v,i}^{(l)}\mathbf{W}^q)(\mathbf{h}_{v,j}^{(l)}\mathbf{W}^k)^{\top}}{\sqrt{d^{(l+1)}}} + \mathbf{b}_{\phi(i,j)}, \quad (2)$$

where $\mathbf{W}^q, \mathbf{W}^k \in \mathbb{R}^{d^{(l)} \times d^{(l+1)}}$ are projection matrices, and the zoom-aware bias $\mathbf{b}_{\phi(i,j)}$ is shared across all layers.

For the propagation process, we feed $\mathbf{H}_{\mathcal{P},v}^{(0)}$ into L stacked Zoom-Aware Transformers, each consisting of a ZMSA model and a feed-forward network (FFN), with layer normalization and residual connections:

$$\begin{aligned} \mathbf{H}'^{(l)}_{\mathcal{P},v} &= \mathrm{ZMSA}(\mathrm{LN}(\mathbf{H}^{(l-1)}_{\mathcal{P},v})) + \mathbf{H}^{(l-1)}_{\mathcal{P},v}, \\ \mathbf{H}^{(l)}_{\mathcal{P},v} &= \mathrm{FFN}(\mathrm{LN}(\mathbf{H}'^{(l)}_{\mathcal{P},v})) + \mathbf{H}'^{(l)}_{\mathcal{P},v}. \end{aligned} \quad (3)$$

After L layers of propagation , we obtain the latent multi-view representation $\mathbf{H}_v^{\mathcal{P}} = \{\mathbf{h}_{v,0}^{\mathcal{P}}, \dots, \mathbf{h}_{v,K}^{\mathcal{P}}\} \in \mathbb{R}^{(K+1)\times d}$.

Dual-View Detach. Next, we construct more informative local (node-level) and global (community-level) representations. The local representation aggregates multi-view tokens based on their relevance to the query node, while the global representation captures structural context reflective of community patterns. The attention weights are computed as:

$$\alpha_i^{loc} = \frac{\exp(\mathbf{h}_{v,0}^{\mathcal{P}} \cdot \mathbf{h}_{v,i}^{\mathcal{P}})}{\sum_{j=0}^{K} \exp(\mathbf{h}_{v,0}^{\mathcal{P}} \cdot \mathbf{h}_{v,j}^{\mathcal{P}})}, \quad (4)$$

$$\alpha_i^{glo} = \frac{\exp\left(\left[\mathbf{h}_{v,0}^{\mathcal{P}} \,\|\, \mathbf{h}_{v,i}^{\mathcal{P}}\right]\mathbf{W}_d\right)}{\sum_{j=1}^{K} \exp\left(\left[\mathbf{h}_{v,0}^{\mathcal{P}} \,\|\, \mathbf{h}_{v,j}^{\mathcal{P}}\right]\mathbf{W}_d\right)}, \quad (5)$$

where $\|$ denotes concatenation, $\mathbf{W}_d \in \mathbb{R}^{2d\times 1}$ is a learnable weight. The local and global representations are obtained as: $\mathbf{h}_{v,loc}^{\mathcal{P}} = \sum_{i=0}^{K} \alpha_i^{loc} \cdot \mathbf{h}_{v,i}^{\mathcal{P}}, \mathbf{h}_{v,glo}^{\mathcal{P}} = \sum_{i=1}^{K} \alpha_i^{glo} \cdot \mathbf{h}_{v,i}^{\mathcal{P}}$.

Attention-based Semantic Aggregation. To effectively integrate semantic features from different meta-paths, we first compute the importance of each meta-path using a shared attention network over the local and global features:

$$\beta_{\mathcal{P}_i} = \frac{\exp(\mathbf{q}_s^\top \cdot \tanh(\mathbf{W}_s(\mathbf{h}_{v,loc}^{\mathcal{P}_i} \,\|\, \mathbf{h}_{v,glo}^{\mathcal{P}_i}) + \mathbf{b}_s))}{\sum_{j=1}^{n} \exp(\mathbf{q}_s^\top \cdot \tanh(\mathbf{W}_s(\mathbf{h}_{v,loc}^{\mathcal{P}_j} \,\|\, \mathbf{h}_{v,glo}^{\mathcal{P}_j}) + \mathbf{b}_s))}, \tag{6}$$

where $\mathbf{W}_s$, $\mathbf{b}_s$, and $\mathbf{q}_s$ are shared parameters across different meta-paths. At last, the final unified local and global representations are then computed as weighted sums, i.e., $\mathbf{h}_v^{loc} = \sum_{i=1}^{n} \beta_{\mathcal{P}_i} \mathbf{h}_{v,loc}^{\mathcal{P}_i}$ and $\mathbf{h}_v^{glo} = \sum_{i=1}^{n} \beta_{\mathcal{P}_i} \mathbf{h}_{v,glo}^{\mathcal{P}_i}$.

Offline Training of VASGhormer

We design a novel semi-supervised strategy, which employs a View-Aware Contrastive Learner and uses node classification as an auxiliary task to enable effective offline training.

View-Aware Contrastive Learner. In real communities, nodes exhibit both structural cohesion and feature similarity. To leverage this property and enable training without community labels, we introduce the View-Aware Contrastive Learner. It integrates a *semantic contrastive loss* from the meta-path view and a *unified contrastive loss* from the unified view, using both classification labels and graph structure for sampling. In this way, the learned representations are more closely align with true community patterns.

Specifically, inspired by the margin-based triplet loss (Schroff, Kalenichenko, and Philbin 2015), we define the *semantic contrastive loss* over meta-path-level representations to achieve discriminative learning from the meta-path view: $\mathcal{L}^{sem} = \frac{1}{|V^t|} \frac{1}{M} \sum_{v \in V^t} \sum_{\mathcal{P} \in \mathcal{P}_s} \frac{1}{|\mathbb{P}_v^{\mathcal{P}}|} \frac{1}{|\mathbb{N}_v^{\mathcal{P}}|} \mathcal{L}_v^{\mathcal{P}}$, where V^t is the set of training nodes, M is the number of meta-paths, and each $\mathcal{L}_v^{\mathcal{P}}$ is computed as follows:

$$\begin{aligned} \mathcal{L}_v^{\mathcal{P}} = \sum_{u \in \mathbb{P}_v^{\mathcal{P}}} \sum_{u' \in \mathbb{N}_v^{\mathcal{P}}} \Big[\max \Big(0, \, & \mathrm{sim}\left(\mathbf{h}_{v,glo}^{\mathcal{P}}, \mathbf{h}_{u',glo}^{\mathcal{P}}\right) \\ -\mathrm{sim}\left(\mathbf{h}_{v,glo}^{\mathcal{P}}, \mathbf{h}_{u,glo}^{\mathcal{P}}\right) - & \mathrm{sim}\left(\mathbf{h}_{v,loc}^{\mathcal{P}}, \mathbf{h}_{v,glo}^{\mathcal{P}}\right) + \epsilon\Big)\Big], \end{aligned} \tag{7}$$

where $\mathrm{sim}(\cdot,\cdot)$ denotes cosine similarity, ϵ is a margin, $\mathbb{P}_v^{\mathcal{P}}$ contains meta-path-based neighbors with the same category label, and $\mathbb{N}_v^{\mathcal{P}}$ comprises nodes at least K-hops away or with different labels, ensuring that sampling respects both structural and attribute-level community patterns. This objective encourages tighter global (community-level) representations for nodes within the same community, separates nodes from different communities, and enforces consistency between the local and global representations of each node.

Additionally, to align with our learning goal under the unified view (after semantic aggregation), we adopt a similar principle and define the *unified contrastive loss* as:

$$\begin{aligned} \mathcal{L}^{uni} = \frac{1}{|V^t|} \frac{1}{|\mathbb{P}_v|} \frac{1}{|\mathbb{N}_v|} \sum_{v \in V^t} \sum_{u \in \mathbb{P}_v} \sum_{u' \in \mathbb{N}_v} \Big[\max \Big(0, \, \mathrm{sim}(\\ \mathbf{h}_v^{glo}, \mathbf{h}_{u'}^{glo}) - \mathrm{sim}\left(\mathbf{h}_v^{glo}, \mathbf{h}_u^{glo}\right) - \mathrm{sim}\left(\mathbf{h}_v^{loc}, \mathbf{h}_v^{glo}\right) + \epsilon\Big)\Big]. \end{aligned} \tag{8}$$

To construct positive and negative samples for $\mathcal{L}^{uni}$, we use a multi-semantic adjacency matrix $\mathbf{A}_{multi}$, where

Algorithm 1: Multi-Constrained Community Search

Input: Query node q, adjacency matrix $\mathbf{A}_{multi}$, community score S_q, score threshold τ, predicted classification $\hat{Y}^{cls}$
Output: Target community C_q

$C_q \leftarrow \{q\}$
$V_{cand} \leftarrow \{v \in V^{tar} \mid S_{q,v} \geq \tau\}$
if $\hat{y}_q^{cls}$ is confident **then**
  $V_{cand} \leftarrow \{v \in V_{cand} \mid \text{argmax } \hat{y}_v^{cls} = \text{argmax } \hat{y}_q^{cls}\}$
end if
$C_q \leftarrow$ largest component in V_{cand} containing q
return C_q

$\mathbf{A}_{multi}(u,v) = 1$ if there exists any $\mathcal{P} \in \mathcal{P}_s$ such that $\mathbf{A}_{u,v}^{\mathcal{P}} = 1$, and 0 otherwise. Then, $\mathbb{P}_v$ contains nodes connected to v in $\mathbf{A}_{multi}$ and sharing the same category label, while $\mathbb{N}_v$ includes nodes that are either unconnected within K-hops or have different labels. Finally, the objective of the View-Aware Contrastive Learner is defined as $\mathcal{L}_{VAC} = \lambda_1 \mathcal{L}^{sem} + \mathcal{L}^{uni}$, where λ_1 is a hyperparameter.

Auxiliary Node Classification. Communities typically consist of densely connected nodes that share the same category, making node classification a valuable prior task for community search. Leveraging this idea, we incorporate a node classification task to enhance the learning of node-level representations $\mathbf{h}_v^{\text{loc}}$. Specifically, we apply a linear projection followed by a sigmoid activation: $\hat{y}_v^{cls} = \sigma(\mathbf{W}_c \cdot \mathbf{h}_v^{loc} + \mathbf{b}_c)$, where $\mathbf{W}_c \in \mathbb{R}^{d_c \times d}$ and $\mathbf{b}_c \in \mathbb{R}^{d_c}$ are trainable parameters. The cross-entropy loss is then adopted as the training loss: $\mathcal{L}_{cls} = -\frac{1}{|V^t|} \sum_{v \in V^t} \sum_{j=1}^{d_c} y_{v,j}^{cls} \log(\hat{y}_{v,j}^{cls})$, where $\hat{y}_v^{cls}$ is the predicted probability and y_v^{cls} is the category label.

Finally, we integrate the two loss functions with a hyperparameter λ_2, and define the overall training objective as $\mathcal{L}_{all} = \mathcal{L}_{VAC} + \lambda_2 \mathcal{L}_{cls}$. The overall training complexity is $O\big(|V^t| \cdot M \cdot L \cdot ((K+1)^2 d + (K+1) d^2)\big)$; additional analysis is provided in the Appendix.

Online Search Strategy

During online search, ML-based methods (Li et al. 2024; Chen et al. 2024) often incur high query costs by relying on model inference to obtain probabilities or scores. To improve efficiency, we precompute and store auxiliary features using the trained VASGhormer, thereby eliminating the need for model propagation during the query phase. Then, we design a Multi-Constrained Community Search approach that incorporates both structural and feature information, and aligns with our auxiliary node classification task.

Specifically, for auxiliary features generation, we employ the trained VASGhormer to obtain community-level representations and classification results for target-type nodes, denoted as $\mathbf{H}^{com} = \{\mathbf{h}_1^{com}, ..., \mathbf{h}_{N_t}^{com}\}$ and $\hat{Y}^{cls} = \{\hat{y}_1^{cls}, ..., \hat{y}_{N_t}^{cls}\}$. Note that $\mathbf{H}^{com}$ and $\hat{Y}^{cls}$ can be reused throughout the online query phase without recomputation.

According to our training objective, nodes with closer community representations are more likely to be tightly connected, share similar features, and belong to the same com-

Method	IMDB			DBLP			ACM			OAG			MAG		
	F1	JAC	NMI	F1	JAC	NMI	F1	JAC	NMI	F1	JAC	NMI	F1	JAC	NMI
ICS-GNN	51.73	42.03	43.47	55.17	43.40	45.06	50.42	41.88	42.12	-	-	-	-	-	-
QD-GNN	53.42	40.92	41.12	58.53	45.82	42.21	53.05	41.03	39.47	-	-	-	-	-	-
COCLEP	57.15	45.19	46.09	61.08	48.05	45.31	58.93	45.19	45.81	35.08	22.02	21.33	-	-	-
TransZero	61.88	48.55	45.76	65.91	50.39	47.73	61.17	49.55	47.22	41.95	24.36	26.89	32.59	24.77	23.83
CS-DAHIN	65.73	50.75	49.58	72.10	55.89	56.03	63.52	53.69	50.46	-	-	-	-	-	-
ST-GNN	66.05	53.67	52.50	77.64	63.03	64.19	70.60	58.76	56.66	45.41	27.79	25.05	-	-	-
FCS-HGNN	65.61	55.24	50.83	79.25	67.97	62.55	68.28	59.34	55.00	44.72	29.15	30.48	-	-	-
SCSAH	**75.34**	**62.80**	**60.29**	**87.76**	**73.12**	**71.44**	**80.71**	**72.09**	**66.50**	**50.42**	**35.92**	**34.26**	**42.45**	**32.98**	**32.53**
Improve (%)	14.62	13.69	14.84	10.74	7.58	11.29	14.32	21.49	17.37	11.03	23.22	12.40	30.25	33.14	36.51

Table 1: Performance comparison with seven baseline methods based on F1, JAC, and NMI (%) across five datasets. Bold numbers indicate the best results, underlined numbers denote the second-best, and "-" indicates out-of-memory.

Dataset	Nodes	Edges	Meta-path
IMDB	*Movie (M): 4,780 Actor (A): 5,841 Director (D): 2,269	M-A: 14,340 M-D: 4,780	MAM MDM
DBLP	*Author (A): 4,057 Paper (P): 14,328 Conference (C): 20 Term (T): 8,789	A-P: 19,645 P-C: 14,328 P-T: 88,420	APA APCPA APTPA
ACM	*Paper (P): 12,499 Author (A): 17,431 Subject (S): 73	P-A: 37,055 P-S: 12,499	PAP PSP
OAG	*Paper (P): 119,483 Author (A): 510,189 Venue (V): 6,934 Institution (I): 9,079	P-A: 340,959 A-A: 329,703 P-V: 119,483 A-I: 612,872	PAP PAIAP
MAG	*Paper (P): 736,389 Author (A): 1,134,649 Field (F): 59,965 Institution (I): 8,740	P-A: 7,145,660 P-P: 5,416,271 P-F: 7,505,078 A-I: 1,043,998	PAP PFP PAIAP

Table 2: Statistics of five real-word datasets.

munity. Thus, for a given query node q, we compute similarity scores to estimate the likelihood that node i is in the same community as q, i.e., $S_{q,i} = \mathrm{sim}(\mathbf{h}_q^{com}, \mathbf{h}_i^{com})$.

Based on the above, we then perform Multi-Constrained Community Search as outlined in Algorithm 1. It first initializes the target community and selects candidate nodes using a score threshold τ (Lines 1-2), where τ can be adjusted based on evaluation results (See Appendix for a Case Study). Next, nodes are conditionally selected if their predicted category are reliable; specifically, a prediction is considered reliable if the gap between the top-1 and top-2 predicted probabilities exceeds the confidence threshold τ_c. This ensures that only high-confident predictions influence community search (Lines 3–4). Finally, given $\mathbf{A}_{multi}$ and V_{cand}, the algorithm returns the largest connected component containing the query node q as the final community (Line 5), ensuring feature coherence and structural connectivity.

Experiments

Experimental Setup

Datasets. Table 2 summarizes the statistics of five real-world AHINs: IMDB (Wang et al. 2019), DBLP (Wang et al. 2019), ACM (Luo et al. 2021), OAG (OAG-L1-Field) (Li et al. 2024) and MAG (OGB-MAG) (Chen et al. 2024; Li et al. 2024) . More details of the datasets and the generation of community labels are provided in the Appendix.

Comparing methods. We compare our approach with existing ML-based CS methods on AHINs, including CS-DAHIN (Song et al. 2024), ST-GNN (Li et al. 2024), and FCS-HGNN (Chen et al. 2024). Due to the limited availability of methods tailored for AHINs, we additionally include several methods designed for homogeneous graphs: ICS-GNN (Gao et al. 2021), QD-GNN (Jiang et al. 2022b), COCLEP (Li et al. 2023), and TransZero (Wang et al. 2024a). Details on how each model is adapted to our datasets and community search task are provided in the Appendix.

Evaluation Metrics. We evaluate performance using three standard metrics based on community labels: F1-score, Jaccard Similarity (JAC), and Normalized Mutual Information (NMI) (Wang et al. 2024a; Chen et al. 2024). Appendix for more detailed environment and parameter settings.

Effectiveness Evaluation

Table 1 shows that SCSAH consistently outperforms all baselines, with average improvements of 16.19% in F1, 19.82% in JAC, and 18.18% in NMI. Graph-dependent models often struggle with scalability and memory usage on large-scale graphs. Among heterogeneous graph baselines, only SCSAH is able to process the MAG dataset without memory issues. Methods designed for homogeneous graphs generally perform poorly due to forced type conversion, which limits their ability to capture essential heterogeneous features. Among AHIN-based methods, CS-DAHIN, originally developed for dynamic graphs, performs inadequately when its dynamic module is removed for static community search. Although ST-GNN and FCS-HGNN show relatively strong baseline performance, they still fall short of SCSAH in both scalability and effectiveness. The full results with mean $\pm$ standard deviation are provided in the Appendix.

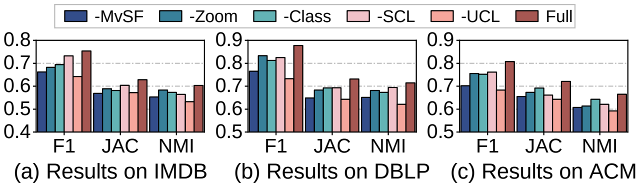

Figure 3: Ablation results of five variants on three datasets.

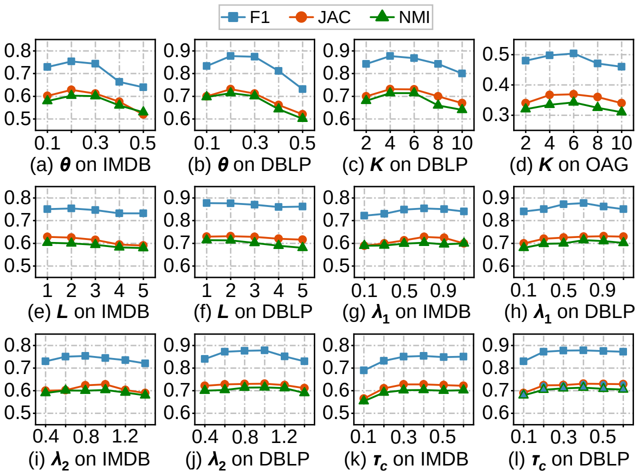

Figure 4: Hyperparameter sensitivity results, with varying cohesion threshold θ (a-b) and hop limitation K (c-d) in MvSF2Token; Transformer layers L (e-f); semantic contrastive loss weight λ_1 (g-h); classification loss weight λ_2 (i-j); and confidence threshold τ_c (k-l) in search Algorithm.

Ablation Study

To evaluate the contribution of each component in SCSAH, we design the following variants: **1) -MvSF**, where the MvSF2Token module is removed and the original attributes of target-type nodes are used, with structural features captured by Hop2Token method (Chen et al. 2023); **2) -Zoom**, which replaces the Zoom-Aware Transformer with a standard Transformer; **3) -Class**, which omits the node classification task from both training and search phases; **4) -SCL**, which eliminates the *semantic contrastive loss*; and **5) -UCL**, which removes the *unified contrastive loss*. As shown in Figure 3, all components have a positive impact on overall performance. Removing the MvSF2Token module results in significant drop, highlighting the critical role of multi-view semantic features in modeling graph structure. The Zoom-Aware Transformer consistently outperforms the standard version, demonstrating its effectiveness in capturing cross-view dependencies. Notably, the *unified contrastive loss* is especially important—its removal leads to an average performance drop of 15.2%—underscoring the central role of unified features in learning effective representations.

Hyperparameter Analysis

We analyze the sensitivity of six key hyperparameters in our model.As shown in Figure 4, the cohesion threshold θ is varied from 0.1 to 0.5, with the best performance at $\theta = 0.2$; higher values filter out too many relevant nodes, thereby hindering multi-view feature construction. The hop limitation K is tested over $\{2, 4, 6, 8, 10\}$, with $K = 4$ performing best on DBLP and $K = 6$ on OAG, indicating that excessively large hop ranges may introduce noise from unrelated nodes. We vary the Transformer layers L from 1 to 5, and observe peak performance at 1 or 2, suggesting that a shallow architecture suffices for capture multi-view interactions. For the semantic contrastive loss weight λ_1, we test values from 0.1 to 1.1, and find $\lambda_1 = 0.7$ achieves the best results, highlighting the greater importance of unified view learning than meta-path view. The classification loss weight λ_2 is also varied from 0.1 to 1.1; performance remains stable between 0.6 and 1.0, while values outside this range cause imbalance between objectives. Finally, the confidence threshold τ_c is tuned from 0.1 to 0.6, with optimal results when τ_c falls between 0.3 and 0.4, effectively filtering out uncertain predictions while preserving reliable classification signals for community search. Overall, most hyperparameters remain stable within a certain range, indicating the robustness of our method across different scenarios without extensive tuning. A more detailed analysis is provided in the Appendix.

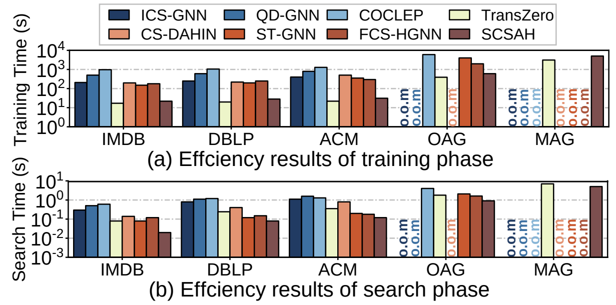

Figure 5: Efficiency evaluation results of training and search phases, where o.o.m denotes out-of-memory.

Efficiency Evaluation

In Figure 5, we report the efficiency results for both training and search phases, evaluated on the same datasets and baseline methods as in the effectiveness evaluation. Note that the training phase begins only after extracting MvSFs for all training nodes, while the querying phase starts after generating the auxiliary features for the full graph. Among all methods, only TransZero and our SCSAH are able train on the large-scale MAG dataset without out-of-memory issues, as both approaches convert structural information into compact tokens, substantially reducing GPU memory overhead during training. Since TransZero is originally designed for homogeneous graphs and does not perform per-meta-path propagation or semantic aggregation, our SCSAH exhibits slightly slower. On datasets where AHIN-specific methods can be executed, SCSAH achieves an average speedup of $10.43\times$ over CS-DAHI, ST-GNN and FCS-HGNN in the training phase, and $4.27\times$ in the search phase.

Conclusion

In this paper, we present SCSAH, a novel framework for community search on AHINs that ensures scalability on large-scale graphs and supports scenarios without community labels. SCSAH comprises three main stages: (1) MvSF2Token, which extracts compact multi-view semantic features as subgraph-level tokens prior to training; (2) offline training of VASGhormer, a view-aware semantic graph transformer, which effectively captures multi-view semantic features and ensures scalability on large graphs, without relying on community labels; and (3) online search, leveraging the generated auxiliary features to perform online community search via a Multi-Constrained Community Search algorithm. Extensive experiments on five real-world datasets demonstrate that SCSAH consistently outperforms existing state-of-the-art methods in both performance and efficiency.

Acknowledgement

This study was funded by the youth talent support program of 'Xing Liao Talent Program' (No. XLYC2203003), the National Natural Science Foundation of China (Nos.62072220, 62472311, 52574191); the Major Program of the National key research and development program (No.2022YFC3004603) and Natural Science Foundation of Liaoning Province (No.2022-KF-13-06, 2022-BS-111); Liaoning Provincial Department of Education Youth Project (No.JYTQN2023189); Natural Science Foundation of Liaoning University (No.LDZDJC2402), Australian Research Council Discovery Early Career Researcher Award (No. DE230100366).

References

Bar-Shalom, G.; Bevilacqua, B.; and Maron, H. 2024. Subgraphormer: Unifying Subgraph GNNs and Graph Transformers via Graph Products. In *Proceedings of the 41st International Conference on Machine Learning*.

Chen, D.; Wang, M.; Chen, H.; Wu, L.; Qin, J.; and Peng, W. 2022. Cross-Modal Retrieval with Heterogeneous Graph Embedding. In *Proceedings of the 30th ACM International Conference on Multimedia*, 3291–3300.

Chen, G.; Guo, F.; Wang, Y.; Liu, Y.; Yu, P.; Shen, H.; and Cheng, X. 2024. FCS-HGNN: Flexible Multi-type Community Search in Heterogeneous Information Networks. In *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*, 207–217. ACM.

Chen, J.; Gao, K.; Li, G.; and He, K. 2023. NAGphormer: A Tokenized Graph Transformer for Node Classification in Large Graphs. In *Proceedings of the 11th International Conference on Learning Representations*.

Fang, Y.; Yang, Y.; Zhang, W.; Lin, X.; and Cao, X. 2020. Effective and Efficient Community Search over Large Heterogeneous Information Networks. In *Proceedings of the VLDB Endowment*, volume 13, 854–867.

Gao, J.; Chen, J.; Li, Z.; and Zhang, J. 2021. ICS-GNN: Lightweight Interactive Community Search via Graph Neural Network. In *Proceedings of the VLDB Endowment*, volume 14, 1006–1018.

Gou, X.; Xu, X.; Wu, X.; Chen, R.; Wang, Y.; Wu, T.; and Ke, X. 2023. Effective and Efficient Community Search with Graph Embeddings. In *Proceedings of the 26th European Conference on Artificial Intelligence*, volume 372, 891–898.

Jiang, Y.; Fang, Y.; Ma, C.; Cao, X.; and Li, C. 2022a. Effective Community Search over Large Star-Schema Heterogeneous Information Networks. In *Proceedings of the VLDB Endowment*, volume 15, 2307–2320.

Jiang, Y.; Rong, Y.; Cheng, H.; Huang, X.; Zhao, K.; and Huang, J. 2022b. Query Driven-Graph Neural Networks for Community Search: From Non-Attributed, Attributed, to Interactive Attributed. In *Proceedings of the VLDB Endowment*, volume 15, 1243–1255.

Li, L.; Luo, S.; Zhao, Y.; Shan, C.; Wang, Z.; and Qin, L. 2023. COCLEP: Contrastive Learning-based Semi-Supervised Community Search. In *Proceedings of the 39th IEEE International Conference on Data Engineering*, 2483–2495.

Li, M.; Cai, X.; Li, L.; Xu, S.; and Ji, H. 2022a. Heterogeneous Graph Attention Network for Drug-Target Interaction Prediction. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, 1166–1176.

Li, X.; Wu, Y.; Ester, M.; Kao, B.; Wang, X.; and Zheng, Y. 2022b. SCHAIN-IRAM: An Efficient and Effective Semi-Supervised Clustering Algorithm for Attributed Heterogeneous Information Networks. *IEEE Transactions on Knowledge and Data Engineering*, 34(4): 1980–1992.

Li, Y.; Chen, X.; Zhao, Y.; Shan, W.; Wang, Z.; Yang, G.; and Wang, G. 2024. Self-Training GNN-based Community Search in Large Attributed Heterogeneous Information Networks. In *Proceedings of the 40th IEEE International Conference on Data Engineering*, 2765–2778.

Liu, Y.; Guo, F.; Xu, B.; Bao, P.; Shen, H.; and Cheng, X. 2024. SACH: Significant-Attributed Community Search in Heterogeneous Information Networks. In *Proceedings of the 40th IEEE International Conference on Data Engineering*, 3283–3296.

Luo, L.; Fang, Y.; Cao, X.; Zhang, X.; and Zhang, W. 2021. Detecting Communities from Heterogeneous Graphs: A Context Path-based Graph Neural Network Model. In *Proceedings of the 30th ACM International Conference on Information and Knowledge Management*, 1170–1180.

Ni, L.; Ye, R.; Luo, W.; Zhang, Y.; Zhang, L.; and Sheng, V. S. 2025. SLRL: Semi-Supervised Local Community Detection Based on Reinforcement Learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 631–639.

Qiao, L.; Zhang, Z.; Yuan, Y.; Chen, C.; and Wang, G. 2021. Keyword-Centric Community Search over Large Heterogeneous Information Networks. In *Proceedings of the 26th International Conference on Database Systems for Advanced Applications*, volume 12681, 158–173.

Rampásek, L.; Galkin, M.; Dwivedi, V. P.; Luu, A. T.; Wolf, G.; and Beaini, D. 2022. Recipe for a General, Powerful, Scalable Graph Transformer. In *Advances in Neural Information Processing Systems*, volume 35, 14501–14515.

Schroff, F.; Kalenichenko, D.; and Philbin, J. 2015. FaceNet: A unified embedding for face recognition and clustering. In *IEEE Conference on Computer Vision and Pattern Recognition*, 815–823.

Shi, C.; Li, Y.; Zhang, J.; Sun, Y.; and Yu, P. S. 2017. A Survey of Heterogeneous Information Network Analysis. *IEEE Transactions on Knowledge and Data Engineering*, 29(1): 17–37.

Song, Y.; Zhou, L.; Yang, P.; Wang, J.; and Wang, L. 2024. CS-DAHIN: Community Search Over Dynamic Attribute Heterogeneous Network. *IEEE Transactions on Knowledge and Data Engineering*, 36(11): 5874–5888.

Wang, J.; Wang, K.; Lin, X.; Zhang, W.; and Zhang, Y. 2024a. Efficient Unsupervised Community Search with Pretrained Graph Transformer. In *Proceedings of the VLDB Endowment*, volume 17, 2227–2240.

Wang, J.; Zhou, L.; Wang, X.; Wang, L.; and Li, S. 2024b. Attribute-sensitive community search over attributed heterogeneous information networks. *Expert Systems with Applications*, 235: 121153.

Wang, X.; Ji, H.; Shi, C.; Wang, B.; Ye, Y.; Cui, P.; and Yu, P. S. 2019. Heterogeneous Graph Attention Network. In *Proceedings of The World Wide Web Conference*, 2022–2032.

Wang, Y.; Gou, X.; Xu, X.; Geng, Y.; Ke, X.; Wu, T.; Yu, Z.; Chen, R.; and Wu, X. 2024c. Scalable Community Search over Large-scale Graphs based on Graph Transformer. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1680–1690.

Wang, Y.; Gu, C.; Xu, X.; Zeng, X.; Ke, X.; and Wu, T. 2024d. Efficient and effective (k, P)-core-based community search over attributed heterogeneous information networks. *Information Sciences*, 661: 120076.

Wang, Z.; Yuan, Y.; Zhou, X.; and Qin, H. 2020. Effective and Efficient Community Search in Directed Graphs Across Heterogeneous Social Networks. In Borovica-Gajic, R.; Qi, J.; and Wang, W., eds., *Proceedings of the 31st Australasian Database Conference on Databases Theory and Applications*, volume 12008, 161–172.

Watts, D. J.; and Strogatz, S. H. 1998. Collective dynamics of 'small-world' networks. *Nature*, 393(6684): 440–442.

Yang, F.; Ma, H.; Gao, W.; and Li, Z. 2022. Community search over heterogeneous information networks via weighting strategy and query replacement. *Frontiers of Computer Science*, 16(4): 164345.

Ye, J.; Zhu, Y.; and Chen, L. 2023. Top-r keyword-based community search in attributed graphs. In *Proceedings of the 39th IEEE International Conference on Data Engineering*, 1652–1664.

Ying, C.; Cai, T.; Luo, S.; Zheng, S.; Ke, G.; He, D.; Shen, Y.; and Liu, T. 2021. Do Transformers Really Perform Badly for Graph Representation? In *Advances in Neural Information Processing Systems*, volume 34, 28877–28888.

Zhong, Q.; Liu, Y.; Ao, X.; Hu, B.; Feng, J.; Tang, J.; and He, Q. 2020. Financial Defaulter Detection on Online Credit Payment via Multi-view Attributed Heterogeneous Information Network. In *Proceedings of The Web Conference*, 785–795.

Zhou, Y.; Fang, Y.; Luo, W.; and Ye, Y. 2023. Influential Community Search over Large Heterogeneous Information Networks. In *Proceedings of the VLDB Endowment*, volume 16, 2047–2060.