CARE: CROSS-DOMAIN ALIGNMENT BY REWEIGHTING EDGES FOR MULTI-VIEW GRAPH DOMAIN ADAPTATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Multi-view graph domain adaptation (MGDA) transfers knowledge from a labeled source graph to an unlabeled target graph, where nodes are connected through multiple relation views. Existing methods typically perform message passing over fixed edge weights, so information specific to a domain or a relation view is aggregated layer after layer along each view's connection pattern and carried into the node representations that alignment must then reconcile. We propose Cross-domain Alignment By Reweighting Edges (CARE), an edge-reweighting framework for MGDA. CARE updates the edge weights of each relation view under the cross-domain alignment objective: edges that propagate information shared across domains are strengthened, and edges that propagate domain-specific or view-specific information are suppressed. Experimental results on a variety of benchmark datasets verify the effectiveness of our method.

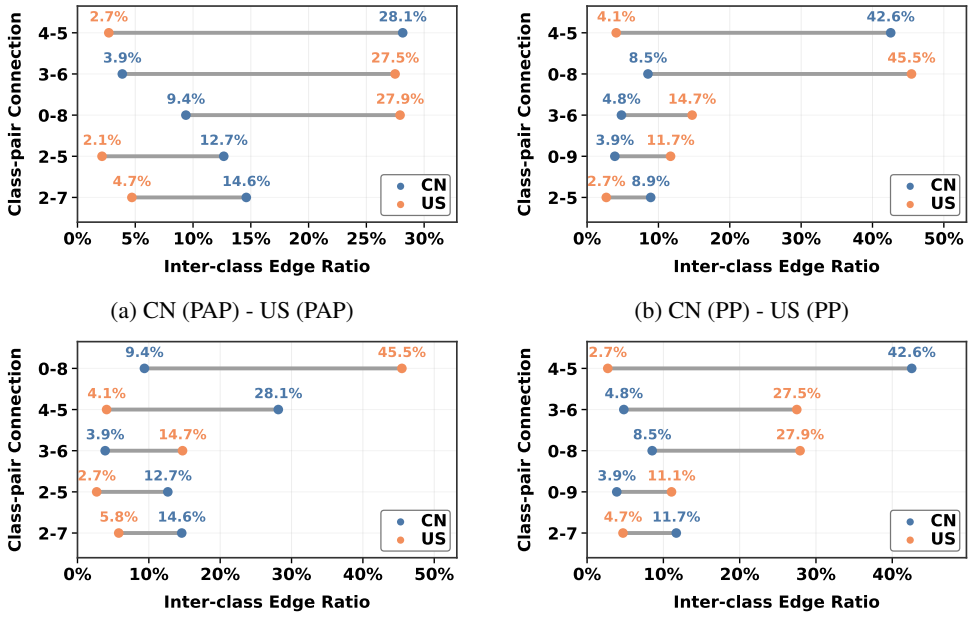
1 INTRODUCTION

Multi-view graphs with a set of nodes and multiple relations can characterize complex correlations among nodes from diverse perspectives, so they are widely used in citation networks, social networks, and knowledge graphs (Wu et al., 2023b; Pan & Kang, 2023). However, multi-view graph datasets often suffer from severe label scarcity due to the high cost and difficulty of annotating complex structured data (Xu et al., 2022; Huang et al., 2024; Zhuo et al., 2026). Therefore, multi-view graph domain adaptation (MGDA) has emerged as an effective paradigm to transfer knowledge from label-rich source domain graphs to unlabeled target domain graphs (Shi et al., 2025a).

As source domain and target domain often differ in node attributes and graph structures, effective cross-domain knowledge transfer remains challenging. Existing methods mitigate domain shifts from multiple perspectives. UDAGCN (Wu et al., 2020) learns domain-invariant representations through adversarial alignment, GAA (Fang et al., 2025a) jointly aligns structural and attribute information across domains. SAGA (Fang et al., 2026) extends adaptation to multi-view graphs by dynamically aligning view-level and hop-level structures.

Despite these advances, existing methods have not explicitly considered the cross-domain differences in edge distributions across different relation views. Specifically, we investigate the cross-domain class-pair edge distributions across different relation views on the real-world datasets. As illustrated in Figure 1, each pair of connected points represents the ratio of a class-pair connection in CN and US graphs in MAG dataset (Shen et al., 2024), with a larger gap indicating a greater cross-domain discrepancy. The observed gaps vary across both same-view and cross-view comparisons of different relation views, indicating that different relation views may exhibit different degrees of domain shift. Under such circumstances, continuing to propagate messages with fixed edge weights may repeatedly aggregate domain-specific and relation-view-specific information along these connection patterns and incorporate this information into node representations, thereby interfering with cross-domain alignment. Therefore, it is necessary to adaptively characterize edge-level cross-domain discrepancies across multiple relation views and adjust the propagation strength of different edges accordingly to improve knowledge transfer.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.



(a) CN (PAP) - US (PAP)

(b) CN (PP) - US (PP)

(c) CN (PAP) - US (PP)

(d) CN (PP) - US (PAP)

Figure 1: Cross-domain discrepancies in class-pair edge distributions across different relation views, including Paper-Author-Paper (PAP) and Paper-Paper (PP). For each class pair, the edge ratio denotes the fraction of inter-class edges connecting that pair, excluding intra-class edges, with longer lines indicating larger cross-domain discrepancies. Additional visualizations of cross-domain edge distributions on other datasets are provided in Appendix F.

To address this issue, we propose Cross-domain Alignment By Reweighting Edges (CARE), an edge reweighting framework for multi-view graph domain adaptation. Rather than performing message passing over fixed edge weights, CARE adaptively updates the edge weights of each relation view under a cross-domain alignment objective. Edges that better support the propagation of cross-domain shared information are regarded as highly transferable edges and assigned larger weights, whereas edges that mainly propagate domain-specific information are regarded as less transferable edges and assigned smaller weights. In this way, cross-domain alignment is explicitly used to guide edge reweighting during message passing, thereby reducing the shift between source domain and target domain.

Our main contributions are summarized as follows:

- We reveal that different relation views exhibit distinct cross-domain edge distribution shifts. This observation suggests that fixed-weight message passing may repeatedly aggregate information specific to a domain or a relation view.
- We propose CARE, an edge reweighting framework for multi-view graph domain adaptation, which utilizes cross-domain representation alignment to guide adaptive edge reweighting across different relation views.
- Comprehensive experiments on benchmarks show the superior performance of our method compared to other state-of-the-art methods for real-world datasets of the cross-network node classification tasks.

2 Related Work

2.1 Multi-View Graph Learning

Multi-view graph learning aims to exploit semantic information from different types of relations and learn more comprehensive node representations. In supervised or semi-supervised settings that rely on label information, SSAMN (Sadikaj et al., 2023) and HGCN-MVSC (Wang

et al., 2024) jointly exploit multiple relation structures, node attributes, and limited labels to learn informative multi-view representations, with the latter further modeling view-specific relations through adaptive graph structure learning. ICGCN (Shi et al., 2025b) alleviates over-smoothing through orthogonality-constrained feature transformation and residual damping, while preserving the expressive power of graph convolution. In contrast, unsupervised multi-view graph learning explores shared and complementary information across views without label supervision. AMTGNN (Shi et al., 2026) captures inter-view correlations, filters low-relevance neighbors, and performs augmented intra-view and inter-view aggregation to improve multi-view graph representation learning. NGCE (Ren et al., 2025) leverages node features to guide multi-view graph encoding, enabling effective integration of homophilic and heterophilic structures while strengthening contrastive learning across views. MGCDC (Wang et al., 2026) integrates view-specific embeddings through view-level attention and jointly enforces cross-view cluster alignment and mutual-information-based semantic consistency to learn a unified consensus representation. Contrastive learning has also become an important paradigm for multi-view graph representation learning. MGCCN (Liu et al., 2022), MGDCR (Mo et al., 2023a), and BTGF (Qian et al., 2024) design different contrastive constraints to promote consistency among multi-view representations and alleviate representation collapse. CoCoMG (Peng et al., 2023) further jointly models consistency and complementarity among multiple graphs through structure preservation and representation correlation constraints. DMG (Mo et al., 2023b) adopts disentangled representation learning to separate multi-view information into common and private components while preserving complementary semantics through contrastive constraints.

2.2 Graph Domain Adaptation

Graph domain adaptation (GDA) aims to transfer knowledge from labeled source graphs to unlabeled target graphs by reducing cross-domain distribution shifts. Early studies mainly focus on learning transferable representations through adversarial or cross-graph alignment (Wu et al., 2020; Guo et al., 2022; Dai et al., 2022). Subsequent methods further investigate graph-specific distribution shifts from the perspectives of graph structure and node attributes, including adaptive shift modeling, propagation adaptation, and spectral alignment (Qiao et al., 2023; Liu et al., 2024a; Yang et al., 2025). GAA (Fang et al., 2025a) jointly aligns structural and attribute information across domains, and highlights the importance of considering both sources of distribution shift for graph adaptation. HGDA (Fang et al., 2025b) further characterizes cross-domain homophily discrepancies and employs mixed graph filters to align homophily-related representations. More recent studies explicitly characterize structural discrepancies through edge or topology adjustment (Liu et al., 2023; 2024b; Wen et al.). ADAlign (Chen et al., 2026a) adaptively models multiple distribution discrepancies and DiffGDA (Chen et al., 2026b) models cross-domain structural-semantic evolution as a continuous diffusion process. DREAM (Yin et al., 2025) combines message passing and path aggregation to capture complementary structural semantics across domains. GMM (Wang et al., 2025) models cross-domain distribution differences with a Gaussian mixture model over graph structures to preserve useful domain-specific information while reducing distribution bias. Beyond conventional GDA settings, recent studies have explored class-imbalanced adaptation (Ma et al., 2025; 2026), information filtering and causal or diffusion-based debiasing (Qiao et al., 2024; Luo et al., 2025; Zhang et al., 2026), source-free adaptation (Luo et al., 2024; Zhang et al., 2024; Luo et al., 2026), as well as open-set and multi-source transfer scenarios (Yin et al., 2024; Zhang & He, 2026; Lin et al., 2025; Shen et al., 2026). More recently, SAGA (Fang et al., 2026) extends GDA to multi-view graphs by dynamically selecting dominant view-hop combinations to guide cross-domain alignment.

3 Method

In this section, we present Cross-domain Alignment By Reweighting Edges (CARE), an edge reweighting framework for cross-domain alignment in multi-view graphs. The overall framework of CARE is shown in Figure 2. Each relation view of source domain and target domain is first reweighted and propagated by APPNP, followed by a shared MLP encoder to obtain view-specific node representations. These representations are aligned through corresponding-view and cross-view objectives and then used for node classification. The entire framework is jointly optimized by the

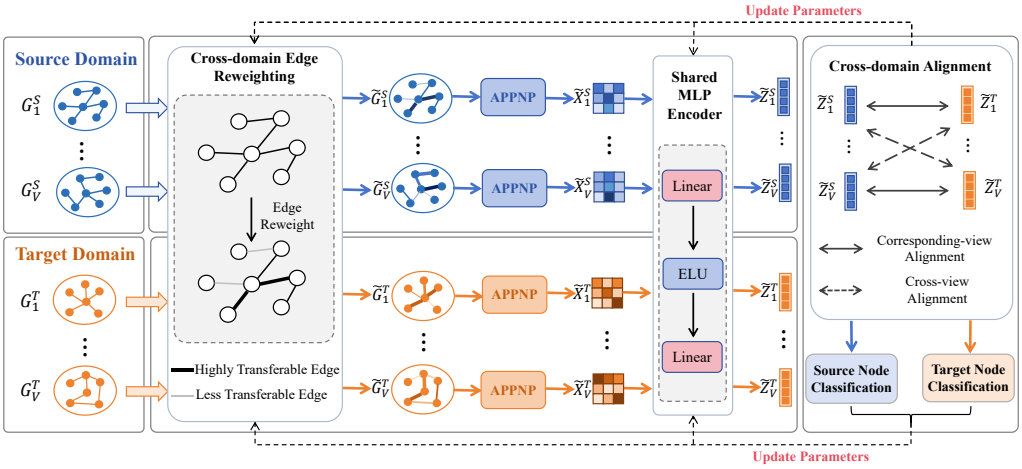


Figure 2: An overview of the proposed CARE framework. CARE adaptively reweights edges in both source domain and target domain under the cross-domain alignment objective, increasing the weights of highly transferable edges while decreasing those of less transferable edges during message propagation.

classification and cross-domain alignment objectives. We describe each component in the following subsections.

3.1 Scalable Multi-view Graph Encoding

We consider an undirected multi-view graph $G = (\mathcal{V}, \{\mathcal{E}_v\}_{v=1}^{V}, \{A_v\}_{v=1}^{V}, X, Y)$, where $\mathcal{V}$ denotes the node set, $\mathcal{E}_v$ and A_v denote the edge set and adjacency matrix of the v-th relation view. X denotes the node feature matrix and Y denotes the label matrix. For domain $D \in \{S, T\}$, the corresponding graph is denoted by G^D. The target-domain G^T labels are unavailable during training.

To efficiently capture the topological information of each view, we decouple graph propagation from feature transformation and employ Approximate Personalized Propagation of Neural Predictions (APPNP) (Gasteiger et al., 2018) to propagate node features. Before propagation, we add self-loops to A_v^D and apply symmetric normalization to obtain the propagation matrix $\bar{A}_v^D$:

$$\bar{A}_v^D = (\bar{\Delta}_v^D)^{-\frac{1}{2}} \left(A_v^D + I\right) (\bar{\Delta}_v^D)^{-\frac{1}{2}} \tag{1}$$

where I is the identity matrix and $\bar{\Delta}_v^D$ denotes the degree matrix of $A_v^D + I$. Then graph propagation is performed on each relation view by APPNP:

$$X_v^{D,(k+1)} = (1-\alpha)\bar{A}_v^D X_v^{D,(k)} + \alpha X^D \tag{2}$$

where $X_v^{D,(k)}$ denotes the node feature matrix after the k-th propagation step in the v-th view of domain D, and $\alpha \in [0, 1]$ is a hyperparameter controlling the contribution of the initial node features during propagation.

After K propagation steps, a shared encoder $f_\Theta(\cdot)$ maps the propagated features into a latent space:

$$Z_v^D = f_\Theta\left(X_v^{D,(K)}\right) \tag{3}$$

where Z_v^D denotes the node representation matrix of the v-th view in domain D. The encoder $f_\Theta(\cdot)$ is implemented as a Multilayer Perceptron (MLP). By separating graph propagation from neural feature transformation, this encoding scheme efficiently aggregates topological information with low computational overhead.

3.2 Edge Reweighting for Cross-Domain Alignment

To improve cross-domain alignment in multi-view graphs, our proposed CARE introduces a learnable edge reweighting module that adaptively adjusts the weights of edges in each relation view

of source domain and target domain. Treating target domain as a fixed reference and reweighting only source domain may cause the model to overfit to target-specific connectivity patterns, rather than focus on information shared across domains, thereby increasing the risk of negative transfer. Therefore, we perform edge reweighting in both domains, allowing the cross-domain alignment objective to jointly optimize message propagation over the source and target views.

For an edge (i, j) in the v-th view of domain D, the edge reweighting module calculates an edge score from its endpoint features and applies Gumbel-Sigmoid reparameterization to obtain a continuous and differentiable reweighting variable:

$$\theta^D_{v,i,j} = \mathrm{MLP}\left([WX^D_i; WX^D_j]\right), \quad \omega^D_{v,i,j} = \mathrm{Sigmoid}\left(\frac{\log\epsilon - \log(1-\epsilon) + \theta^D_{v,i,j}}{\tau_{\mathrm{aug}}}\right) \tag{4}$$

where X^D_i and X^D_j denote the feature vectors of nodes i and j, respectively. W denotes the feature transformation parameters and $[\cdot;\cdot]$ denotes feature concatenation. $\theta^D_{v,i,j}$ and $\omega^D_{v,i,j}$ denote the edge score and continuous reweighting variable respectively. $\epsilon \sim \mathrm{Uniform}(0, 1)$ is independently sampled, and τ_{aug} denotes the temperature parameter.

To keep the edge propagation strength within a reasonable range, we center the reweighting variables within each relation view. Specifically, their mean is computed as

$$\mu^D_v = \frac{1}{|\mathcal{E}^D_v|} \sum_{(i,j)\in\mathcal{E}^D_v} \omega^D_{v,i,j}. \tag{5}$$

Each centered variable is normalized by the maximum deviation from the mean within the same relation view and mapped to a bounded reweighting factor:

$$\tilde{\omega}^D_{v,i,j} = 1 + \rho \frac{\omega^D_{v,i,j} - \mu^D_v}{\max\left(\max_{(m,n)\in\mathcal{E}^D_v} \left|\omega^D_{v,m,n} - \mu^D_v\right|, \xi\right)}. \tag{6}$$

where (m, n) indexes the edges in $\mathcal{E}^D_v$ when calculating the maximum deviation. The inner maximum selects the largest absolute deviation from the mean among all edges in the relation view, while the outer maximum with $\xi > 0$ prevents the denominator from becoming too small. The parameter $\rho \in [0, 1]$ controls the maximum reweighting magnitude, while $\xi > 0$ is a small constant for numerical stability. This formulation constrains $\tilde{\omega}^D_{v,i,j} \in [1-\rho, 1+\rho]$.

The resulting bounded reweighting factor is applied to each existing edge:

$$\tilde{A}^D_{v,i,j} = A^D_{v,i,j}\tilde{\omega}^D_{v,i,j}. \tag{7}$$

where $\tilde{A}^D_v$ denotes the reweighted adjacency matrix. The resulting reweighted graph is defined as $\tilde{G}^D_v = (\mathcal{V}^D, \mathcal{E}^D_v, \tilde{A}^D_v, X^D)$, where the weights of existing edges are adaptively adjusted.

Following the same normalization, propagation, and encoding procedures described in the previous subsection, the propagated node features $\tilde{X}^{D,(K)}_v$ and the corresponding node representations $\tilde{Z}^D_v$ are obtained from the reweighted graph $\tilde{G}^D_v$.

To optimize edge reweighting for cross-domain alignment, an alignment objective is defined in the representation space. Specifically, for the u-th source view and the v-th target view, their node representations are mapped into an alignment space through view-specific projection heads:

$$P^S_u = g_u(\tilde{Z}^S_u), \qquad P^T_v = g_v(\tilde{Z}^T_v) \tag{8}$$

where $g_u(\cdot)$ and $g_v(\cdot)$ denote the projection heads associated with the u-th and v-th relation views respectively. P^S_u and P^T_v denote the projected representation matrices of the u-th source view and the v-th target view.

For each source node i in the current batch, its cosine similarities are calculated with all target nodes. The source-to-target alignment loss is defined as:

$$\mathcal{L}^{u,v}_{S\to T} = -\frac{1}{N_S}\sum_{i=1}^{N_S} \log\left[\sum_{j=1}^{N_T} \exp\left(\frac{s(p^S_{u,i}, p^T_{v,j})}{\tau}\right)\right] + \log(N_S N_T) \tag{9}$$

where $p^S_{u,i}$ and $p^T_{v,j}$ denote the projected representations of source node i and target node j in P^S_u and P^T_v. $s(\cdot,\cdot)$ denotes cosine similarity and τ is the temperature parameter, N_S and N_T denote the numbers of source nodes and target nodes in the current batch. By aggregating cross-domain similarities, this objective encourages stronger representation correlations between source domain and target domain.

Similarly, the target-to-source alignment loss $\mathcal{L}^{v,u}_{T\to S}$ is calculated. Combining the two directions, the bidirectional alignment loss for a view pair (u, v) is defined as:

$$\mathcal{L}^{u,v}_{\mathrm{C}} = \mathcal{L}^{u,v}_{S\to T} + \mathcal{L}^{v,u}_{T\to S}. \tag{10}$$

To exploit cross-domain information across multiple relation views, we jointly consider corresponding-view and cross-view alignment. The optimization objective for edge reweighting is defined as:

$$\mathcal{L}_{\mathrm{A}} = \frac{1}{V}\sum_{v=1}^{V}\mathcal{L}^{v,v}_{\mathrm{C}} + \lambda\frac{1}{V(V-1)}\sum_{\substack{u,v=1\\ u\neq v}}^{V}\mathcal{L}^{u,v}_{\mathrm{C}} \tag{11}$$

where the first term aligns corresponding relation views across domains, while the second captures cross-domain associations between different relation views. λ controls the contribution of cross-view alignment.

Since the node representations $\tilde{Z}^D_v$ are obtained through propagation over the reweighted graphs $\tilde{G}^D_v$, gradients from $\mathcal{L}_{\mathrm{A}}$ can be backpropagated through the encoder and graph propagation process to the edge reweighting module, thereby updating the reweighting factors $\omega^D_{v,i,j}$. In this way, CARE adaptively adjusts the propagation strength of individual edges under the cross-domain alignment objective, yielding node representations $\tilde{Z}^D_v$ that better support cross-domain alignment.

3.3 Node Classification and Joint Optimization

For labeled nodes in source domain, the node classification is performed based on the multi-view representations learned from the reweighted graphs, and the model is optimized by cross-entropy loss. The source classification loss is defined as:

$$\mathcal{L}_S = -\frac{1}{N_S}\sum_{i=1}^{N_S}\sum_{c=1}^{C} y^S_{i,c}\log\hat{y}^S_{i,c} \tag{12}$$

where C is the number of classes, $y^S_{i,c}$ denotes the ground-truth label of the i-th source node for class c, and $\hat{y}^S_{i,c}$ is the corresponding predicted probability.

Since target labels are unavailable during training, the unlabeled target data is further exploited through prediction entropy minimization. The target-domain loss is defined as:

$$\mathcal{L}_T = -\frac{1}{N_T}\sum_{i=1}^{N_T}\sum_{c=1}^{C} \hat{y}^T_{i,c}\log\hat{y}^T_{i,c} \tag{13}$$

where $\hat{y}^T_{i,c}$ denotes the predicted probability that the i-th target node belongs to class c.

Combining the cross-domain alignment objective with the source classification loss and target entropy loss, the overall objective is defined as:

$$\mathcal{L} = \mathcal{L}_S + \delta\mathcal{L}_T + \gamma\mathcal{L}_{\mathrm{A}} \tag{14}$$

where δ and γ are hyperparameters that control the contributions of the corresponding loss terms.

4 EXPERIMENT

4.1 Experimental Settings

Downstream Tasks We focus on the node classification task to evaluate the effectiveness of the proposed CARE framework for multi-view graph domain adaptation. In this setting, source

Table 1: Statistics of experimental datasets.

Graph	Nodes	Relation views	Edges per view	Feature dimension	Classes
ACM1	3,025	PAP / PSP	26,256 / 2,207,736	512	3
ACM2	4,019	PAP / PSP	57,853 / 4,338,213	512	3
CN	25,467	PAP / PP	157,303 / 43,683	128	10
US	55,392	PAP / PP	12,332,742 / 289,833	128	10
JP	12,651	PAP / PP	41,165 / 20,747	128	10
FR	12,052	PAP / PP	51,739 / 30,915	128	10
DE	17,919	PAP / PP	77,100 / 45,528	128	10
RU	3,690	PAP / PP	3,315 / 3,539	128	10

domain provides labeled nodes for model training, while target domain only contains unlabeled nodes during the adaptation process. The objective is to transfer transferable knowledge from source domain to target domain and improve classification performance under domain distribution shifts. Following previous studies on multi-view graph domain adaptation, classification accuracy (ACC) and Macro-F1 are utilized as evaluation metrics.

Datasets To validate the effectiveness of our proposed CARE framework on the domain adaptation node classification task, we utilize real-world citation network datasets in our empirical study, including ACM (Shen et al., 2024) and MAG (Shen et al., 2024). ACM1 and ACM2 derived from ACM contain two relation views of Paper-Author-Paper (**PAP**) and Paper-Subject-Paper (**PSP**), with node labels covering three research areas. MAG consists of six country-specific subgraphs, including China (**CN**), the United States (**US**), Germany (**DE**), France (**FR**), Japan (**JP**), and Russia (**RU**), and each subgraph contains two relation views of PAP and Paper-Paper (**PP**), with node labels covering ten research topics. The basic statistics of these datasets, including the number of nodes, relation views, edges per view, feature dimensions, and classes, are summarized in Table 1. Additional structural statistics are provided in Appendix A. We evaluate two bidirectional transfer tasks on ACM and ten transfer tasks on MAG. The specific transfer directions are listed in Table 2 and Table 3.

Baselines We compare our proposed CARE with several representative baselines for multi-view graph domain adaptation. **GCN** (Kipf & Welling, 2016) and **GAT** (Veličković et al., 2017) serve as non-adaptive baselines, using standard graph convolution and attention-based neighbor aggregation, respectively. **AdaGCN** (Dai et al., 2022) adaptively adjusts graph propagation based on node-level or edge-level information. Among graph domain adaptation methods, **UDAGCN** (Wu et al., 2020) employs dual graph convolutional networks with adversarial domain alignment, while **GRADE-N** (Wu et al., 2023a) measures structural shifts using graph subtree discrepancy.

Table 2: Cross-network node classification on ACM datasets.

Methods	ACM1→ACM2		ACM2→ACM1	
	ACC	F1	ACC	F1
GCN	0.367	0.253	0.359	0.236
GAT	0.373	0.319	0.363	0.241
AdaGCN	0.431	0.272	0.386	0.362
UDAGCN	0.452	0.283	0.409	0.347
GRADE-N	0.469	0.291	0.373	0.343
ACDNE	0.457	0.261	0.392	0.363
ASN	0.436	0.321	0.437	0.327
SpecReg	0.493	0.308	0.431	0.381
PA	0.506	0.333	0.452	0.377
GraphATA	0.501	0.322	0.401	0.395
HGDA	0.511	0.311	0.441	0.414
SAGA	<u>0.523</u>	<u>0.351</u>	<u>0.454</u>	<u>0.444</u>
CARE	**0.538**	**0.432**	**0.541**	**0.535**

ACDNE (Shen et al., 2020) performs cross-network feature alignment. **ASN** (Zhang et al., 2021) learns domain-invariant representations through shared and private encoders. **SpecReg** (You et al., 2023) improves cross-domain generalization through spectral regularization, whereas **PA** (Liu et al., 2024b) recalibrates edge weights to alleviate structural and label shifts. **GraphATA** (Zhang & He, 2025) performs node-wise personalized aggregation of source models, and **HGDA** (Fang et al., 2025b) aligns homophily-related and attribute information across domains using specialized graph filters. **SAGA** (Fang et al., 2026) dynamically selects dominant view-hop combinations and performs cross-domain alignment to reduce multi-view structural shifts.

4.2 Results and Analyses

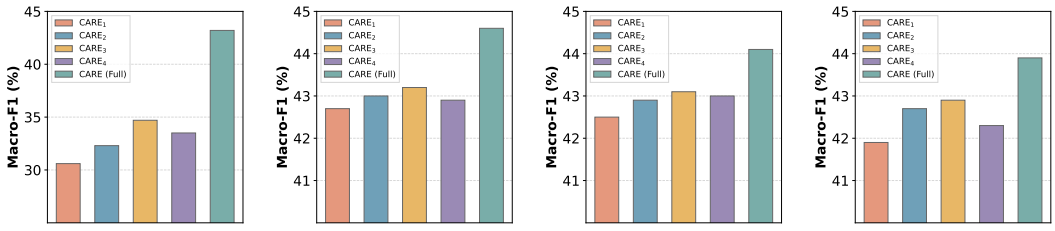
The experimental results are summarized in Table 2 and Table 3, where the best performance is highlighted in bold. Overall, our proposed CARE achieves the highest ACC and Macro-F1 scores across all 12 source-target domain pairs. Compared with the strongest baseline SAGA, CARE improves the ACC on ACM2→ACM1 from 45.4% to 54.1%, and the Macro-F1 from 44.4% to 53.5%. On the ten MAG transfer tasks, the average ACC increases from 48.25% to 49.96%, while the average Macro-F1 increases from 38.26% to 41.06%. This illustrates that adaptively reweighting edges under the cross-domain alignment objective can better capture information shared between the source domain and target domain, thereby improving transfer performance. To further evaluate the quality of the learned representations, we also compare CARE with representative baseline methods using t-SNE, with detailed results provided in Appendix D.

Table 3: Cross-network node classification on MAG datasets.

Methods	Metric	CN→US	US→CN	JP→CN	CN→JP	DE→FR	FR→DE	RU→US	US→RU	DE→CN	CN→DE
GCN	ACC	0.351	0.263	0.327	0.335	0.319	0.302	0.301	0.317	0.281	0.291
	Macro-F1	0.217	0.218	0.221	0.201	0.203	0.167	0.233	0.320	0.167	0.223
GAT	ACC	0.325	0.316	0.361	0.309	0.324	0.253	0.265	0.339	0.281	0.316
	Macro-F1	0.218	0.320	0.253	0.237	0.212	0.249	0.221	0.323	0.249	0.209
AdaGCN	ACC	0.418	0.483	0.389	0.425	0.431	0.314	0.370	0.435	0.359	0.363
	Macro-F1	0.231	0.362	0.257	0.244	0.242	0.285	0.242	0.369	0.265	0.212
UDAGCN	ACC	0.425	0.453	0.451	0.414	0.352	0.364	0.359	0.421	0.356	0.369
	Macro-F1	0.243	0.341	0.271	0.275	0.274	0.262	0.239	0.361	0.296	0.253
GRADE-N	ACC	0.421	0.467	0.376	0.404	0.455	0.375	0.353	0.429	0.358	0.362
	Macro-F1	0.241	0.352	0.305	0.292	0.286	0.287	0.245	0.347	0.299	0.252
ACDNE	ACC	0.410	0.458	0.422	0.436	0.463	0.359	0.351	0.434	0.360	0.378
	Macro-F1	0.225	0.352	0.335	0.275	0.279	0.292	0.231	0.388	0.265	0.291
ASN	ACC	0.416	0.431	0.335	0.432	0.433	0.368	0.289	0.435	0.334	0.372
	Macro-F1	0.232	0.387	0.219	0.251	0.248	0.297	0.216	0.226	0.258	0.270
SpecReg	ACC	0.425	0.501	0.363	0.419	0.424	0.387	0.328	0.429	0.330	0.371
	Macro-F1	0.247	0.380	0.273	0.263	0.236	0.282	0.213	0.371	0.279	0.303
PA	ACC	0.402	0.529	0.351	0.323	0.416	0.357	0.306	0.317	0.336	0.357
	Macro-F1	0.236	0.331	0.295	0.243	0.229	0.299	0.235	0.367	0.286	0.263
GraphATA	ACC	0.407	0.518	0.382	0.341	0.327	0.414	0.351	0.341	0.341	0.321
	Macro-F1	0.241	0.339	0.287	0.252	0.291	0.289	0.234	0.351	0.274	0.254
HGDA	ACC	0.417	0.529	0.392	0.343	0.391	0.421	0.361	0.337	0.361	0.317
	Macro-F1	0.239	0.333	0.321	0.251	0.284	0.279	0.212	0.344	0.271	0.249
SAGA	ACC	0.553	0.535	0.512	0.557	0.529	0.515	0.377	0.441	0.425	0.381
	Macro-F1	0.341	0.426	0.372	0.415	0.426	0.401	0.378	0.397	0.320	0.350
CARE	ACC	**0.593**	**0.539**	**0.532**	**0.577**	**0.543**	**0.541**	**0.389**	**0.449**	**0.444**	**0.389**
	Macro-F1	**0.420**	**0.439**	**0.402**	**0.446**	**0.441**	**0.451**	**0.387**	**0.405**	**0.354**	**0.361**

4.3 Ablation Study

To validate the effectiveness of different components in CARE, we compare the complete model with its four variants on the ACM and MAG datasets. Additional ablation results on the remaining transfer tasks are provided in Appendix G.



(a) ACM1 → ACM2. (b) CN → JP. (c) DE → FR. (d) US → CN.

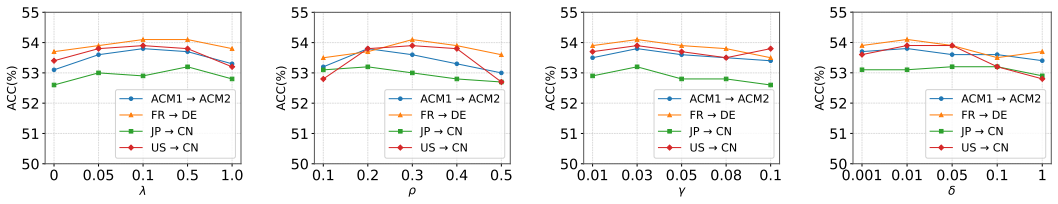
Figure 3: Classification Macro-F1 of different variants of CARE on the ACM and MAG datasets.

- **CARE$_1$:** Removes edge reweighting and uses the original edge weights. This variant evaluates the effectiveness of adaptive edge reweighting.
- **CARE$_2$:** Removes the cross-domain alignment guidance for edge reweighting by stopping the alignment gradients to the edge-reweighting module, while retaining cross-domain alignment for representation learning.
- **CARE$_3$:** Removes cross-view alignment. This variant evaluates the benefit of exploiting complementary information across relation views.
- **CARE$_4$:** Removes edge reweighting on the target domain and only reweights the source graph. This variant evaluates the effectiveness of joint edge reweighting in both domains.
- **CARE (Full):** The complete model with all components.

According to Figure 3, we can draw the following conclusions: (1) CARE consistently outperforms all its variants, demonstrating the effectiveness of the complete framework. (2) Adaptive edge reweighting is important for adjusting information propagation under domain shifts. (3) Cross-view alignment and joint edge reweighting in both source domain and target domain improve performance by exploiting complementary information across relation views and facilitating cross-domain alignment.

4.4 Hyperparameter Sensitivity Analysis

We analyze the sensitivity of CARE to four hyperparameters: the cross-view alignment weight λ, the edge reweighting strength ρ, the alignment weight γ, and the target entropy loss weight δ, as shown in Figure 4. Additional parameter analyses are provided in Appendix H. CARE shows relatively low sensitivity to λ, with only small performance variations across different settings. For ρ, moderate values generally achieve better results, whereas further increasing ρ leads to performance degradation, possibly because excessive edge reweighting disturbs the original information propagation. CARE also performs better with a relatively small γ, while a large alignment weight may overemphasize cross-domain alignment and weaken discriminative information. Similar behavior can be observed for δ, where small values generally yield better performance. Overall, although the optimal values vary across transfer tasks, CARE remains stable within a reasonable parameter range.



(a) The influence of λ. (b) The influence of ρ. (c) The influence of γ. (d) The influence of δ.

Figure 4: Hyperparameter sensitivity analysis of CARE on different transfer tasks.

5 Conclusion

In this paper, we propose CARE to address the differentiated cross-domain edge distribution shifts across different relation views in multi-view graph domain adaptation. We reveal that different relation views exhibit distinct edge distribution discrepancies between source domain and target domain, making it difficult for fixed edge weights to adapt information propagation to their transfer characteristics. To address this issue, CARE utilizes cross-domain alignment signals to adaptively reweight individual edges across different views in both source domain and target domain. Extensive experiments on real-world multi-relational graph benchmarks demonstrate that CARE achieves competitive performance in cross-domain node classification, while ablation and parameter studies further verify the effectiveness of its key designs. Future work will explore more expressive and adaptive strategies for characterizing fine-grained cross-domain edge distribution shifts and incorporating them into multi-view graph representation learning.

6 AI USE STATEMENT

In this work, we used generative AI tools solely for language polishing. We have not used generative AI tools for literature retrieval or discovery, research ideation or execution, drafting sections of the paper, generating synthetic datasets, or proving mathematical claims. We have reviewed all AI-assisted revisions for accuracy, clarity, and consistency with the original technical content. We take responsibility for the final content of this work, including text, claims, or artifacts produced with the aid of generative AI.

REFERENCES

Wei Chen, Xingyu Guo, Shuang Li, Zhao Zhang, Yan Zhong, Fuzhen Zhuang, and deqing wang. Learning adaptive distribution alignment with neural characteristic function for graph domain adaptation. In C. Vondrick, B. Hariharan, C. Raffel, L. Pinto, D. Yang, and A. Faust (eds.), *International Conference on Learning Representations*, volume 2026, pp. 17271–17294, 2026a.

Wei Chen, Xingyu Guo, Shuang Li, Yan Zhong, Zhao Zhang, Fuzhen Zhuang, Hongrui Liu, Libang Zhang, Guo Ye, and Huimei He. Learning structure-semantic evolution trajectories for graph domain adaptation. *International Conference on Learning Representations*, 2026b.

Quanyu Dai, Xiao-Ming Wu, Jiaren Xiao, Xiao Shen, and Dan Wang. Graph transfer learning via adversarial domain adaptation with graph convolution. *IEEE Transactions on Knowledge and Data Engineering*, 35(5):4908–4922, 2022.

Ruiyi Fang, Bingheng Li, Qiuhao Zeng, Nima Hosseini Dashtbayaz, Ruizhi Pu, Charles Ling, Boyu Wang, et al. On the benefits of attribute-driven graph domain adaptation. In *International conference on learning representations*, volume 2025, pp. 30762–30782, 2025a.

Ruiyi Fang, Bingheng Li, Jingyu Zhao, Ruizhi Pu, Qiuhao Zeng, Gezheng Xu, Charles Ling, and Boyu Wang. Homophily enhanced graph domain adaptation. *arXiv preprint arXiv:2505.20089*, 2025b.

Ruiyi Fang, Jingyu Zhao, Shuo Wang, Ruizhi Pu, Bingheng Li, Jiale Cai, Zhihao Li, Zihao Jing, Jian Zhu, Song Tang, et al. Saga: Structural aggregation guided alignment with dynamic view and neighborhood order selection for multiview graph domain adaptation. In *International Conference on Learning Representations*, volume 2026, pp. 95585–95607, 2026.

Johannes Gasteiger, Aleksandar Bojchevski, and Stephan Günnemann. Predict then propagate: Graph neural networks meet personalized pagerank. *arXiv preprint arXiv:1810.05997*, 2018.

Gaoyang Guo, Chaokun Wang, Bencheng Yan, Yunkai Lou, Hao Feng, Junchao Zhu, Jun Chen, Fei He, and Philip S Yu. Learning adaptive node embeddings across graphs. *IEEE Transactions on Knowledge and Data Engineering*, 35(6):6028–6042, 2022.

Jincheng Huang, Yujie Mo, Xiaoshuang Shi, Lei Feng, and Xiaofeng Zhu. Enhancing the influence of labels on unlabeled nodes in graph convolutional networks. *arXiv preprint arXiv:2411.02279*, 2024.

Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.

Chuanyun Lin, Xi Zhou, and Xiao Shen. Hierarchical graph contrastive domain adaptation for multi-source cross-network node classification. *Expert Systems with Applications*, 284:127900, 2025.

Liang Liu, Zhao Kang, Jiajia Ruan, and Xixu He. Multilayer graph contrastive clustering network. *Information Sciences*, 613:256–267, 2022.

Meihan Liu, Zeyu Fang, Zhen Zhang, Ming Gu, Sheng Zhou, Xin Wang, and Jiajun Bu. Rethinking propagation for unsupervised graph domain adaptation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 13963–13971, 2024a.

Shikun Liu, Tianchun Li, Yongbin Feng, Nhan Tran, Han Zhao, Qiang Qiu, and Pan Li. Structural re-weighting improves graph domain adaptation. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett (eds.), *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 21778–21793. PMLR, 23–29 Jul 2023.

Shikun Liu, Deyu Zou, Han Zhao, and Pan Li. Pairwise alignment improves graph domain adaptation. In Ruslan Salakhutdinov, Zico Kolter, Katherine Heller, Adrian Weller, Nuria Oliver, Jonathan Scarlett, and Felix Berkenkamp (eds.), *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 32552–32575. PMLR, 21–27 Jul 2024b.

Junyu Luo, Zhiping Xiao, Yifan Wang, Xiao Luo, Jingyang Yuan, Wei Ju, Langechuan Liu, and Ming Zhang. Rank and align: Towards effective source-free graph domain adaptation. In Kate Larson (ed.), *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI-24*, pp. 4706–4714. International Joint Conferences on Artificial Intelligence Organization, 8 2024. doi: 10.24963/ijcai.2024/520. Main Track.

Junyu Luo, Yuhao Tang, Yiwei Fu, Xiao Luo, Zhizhuo KOU, Zhiping Xiao, Wei Ju, Wentao Zhang, and Ming Zhang. Sparse causal discovery with generative intervention for unsupervised graph domain adaptation. In *Forty-second International Conference on Machine Learning*, 2025.

Junyu Luo, Haoyu Tao, Xiao Luo, Yusheng Zhao, Zhiping Xiao, Dailan He, Wei Ju, Chong Chen, Xian-Sheng Hua, and Ming Zhang. Robust cross supervision with target mining for source-free graph domain adaptation. *IEEE Transactions on Knowledge and Data Engineering*, 2026.

Xin Ma, Yifan Wang, Siyu Yi, Wei Ju, Bei Wu, Ziyue Qiao, Chenwei Tang, and Jiancheng Lv. Pala: Class-imbalanced graph domain adaptation via prototype-anchored learning and alignment. In *IJCAI*, pp. 3198–3207, 2025.

Xin Ma, Yifan Wang, Siyu Yi, Wei Ju, Junyu Luo, Yusheng Zhao, Xiao Luo, and Jiancheng Lv. Dual prototype-enhanced contrastive framework for class-imbalanced graph domain adaptation. *Advances in Neural Information Processing Systems*, 38:48456–48482, 2026.

Yujie Mo, Yuhuan Chen, Yajie Lei, Liang Peng, Xiaoshuang Shi, Changan Yuan, and Xiaofeng Zhu. Multiplex graph representation learning via dual correlation reduction. *IEEE Transactions on Knowledge and Data Engineering*, 35(12):12814–12827, 2023a.

Yujie Mo, Yajie Lei, Jialie Shen, Xiaoshuang Shi, Heng Tao Shen, and Xiaofeng Zhu. Disentangled multiplex graph representation learning. In *International conference on machine learning*, pp. 24983–25005. PMLR, 2023b.

Erlin Pan and Zhao Kang. High-order multi-view clustering for generic data. *Information Fusion*, 100:101947, 2023.

Liang Peng, Xin Wang, and Xiaofeng Zhu. Unsupervised multiplex graph learning with complementary and consistent information. In *Proceedings of the 31st ACM International Conference on Multimedia*, pp. 454–462, 2023.

Xiaowei Qian, Bingheng Li, and Zhao Kang. Upper bounding barlow twins: A novel filter for multi-relational clustering. In *Proceedings of the AAAI conference on artificial intelligence*, volume 38, pp. 14660–14668, 2024.

Ziyue Qiao, Xiao Luo, Meng Xiao, Hao Dong, Yuanchun Zhou, and Hui Xiong. Semi-supervised domain adaptation in graph transfer learning. *arXiv preprint arXiv:2309.10773*, 2023.

Ziyue Qiao, Meng Xiao, Weiyu Guo, Xiao Luo, and Hui Xiong. Information filtering and interpolating for semi-supervised graph domain adaptation. *Pattern Recognition*, 153:110498, 2024.

Yazhou Ren, JunLong Ke, Zichen Wen, Tianyi Wu, Yang Yang, Xiaorong Pu, and Lifang He. Multi-view graph clustering via node-guided contrastive encoding. In *ICML*, 2025.

Ylli Sadikaj, Justus Rass, Yllka Velaj, and Claudia Plant. Semi-supervised embedding of attributed multiplex networks. In *Proceedings of the ACM Web Conference 2023*, pp. 578–587, 2023.

Xiao Shen, Quanyu Dai, Fu-lai Chung, Wei Lu, and Kup-Sze Choi. Adversarial deep network embedding for cross-network node classification. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pp. 2991–2999, 2020.

Xiao Shen, Chuanyun Lin, Zhihao Chen, Junwei Gong, and Xi Zhou. Prototypical multi-source graph contrastive domain adaptation. *Pattern Recognition*, pp. 114312, 2026.

Zhixiang Shen, Haolan He, and Zhao Kang. Balanced multi-relational graph clustering. In *Proceedings of the 32nd ACM international conference on multimedia*, pp. 4120–4128, 2024.

Bo-Shen Shi, Yong-Qing Wang, Fang-Da Guo, Bing-Bing Xu, Hua-Wei Shen, and Xue-Qi Cheng. Domain adaptation for graph representation learning: Challenges, progress, and prospects. *Journal of Computer Science and Technology*, 40(2):283–300, 2025a.

Long Shi, Junyu Chen, Lei Cao, Jun Wang, Jinghua Tan, and Badong Chen. Multi-view graph neural networks by augmented aggregation. *Applied Soft Computing*, pp. 114621, 2026.

Yongquan Shi, Yueyang Pi, Zhanghui Liu, Hong Zhao, and Shiping Wang. Information-controlled graph convolutional network for multi-view semi-supervised classification. *Neural Networks*, 184:107102, 2025b.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. *arXiv preprint arXiv:1710.10903*, 2017.

Mengzhu Wang, Wenhao Ren, Yu Zhang, Yanlong Fan, Dianxi Shi, Luoxi Jing, and Nan Yin. Gaussian mixture model for graph domain adaptation. In *IJCAI*, pp. 1963–1972, 2025.

Shiping Wang, Sujia Huang, Zhihao Wu, Rui Liu, Yong Chen, and Dell Zhang. Heterogeneous graph convolutional network for multi-view semi-supervised classification. *Neural Networks*, 178:106438, 2024.

Zuowei Wang, Sen Xu, Naixuan Guo, Xuesheng Bian, Xiufang Xu, Shanliang Yao, Xianye Ben, and Tian Zhou. Multi-view graph clustering via dual attention fusion and collaborative optimization. *Neural Networks*, pp. 108704, 2026.

Hongwei Wen, Can Zhang, Haoyu He, Hanyuan Hang, and Minglong Lei. Progressive graph structure adjustment for homophily shift adaptation. In *Forty-third International Conference on Machine Learning*.

Jun Wu, Jingrui He, and Elizabeth Ainsworth. Non-iid transfer learning on graphs. In *Proceedings of the AAAI conference on artificial intelligence*, volume 37, pp. 10342–10350, 2023a.

Lingfei Wu, Yu Chen, Kai Shen, Xiaojie Guo, Hanning Gao, Shucheng Li, Jian Pei, and Bo Long. Graph neural networks for natural language processing: A survey. *Foundations and Trends in Machine Learning*, 16(2):119–328, 2023b.

Man Wu, Shirui Pan, Chuan Zhou, Xiaojun Chang, and Xingquan Zhu. Unsupervised domain adaptive graph convolutional networks. In *Proceedings of the web conference 2020*, pp. 1457–1467, 2020.

Zihao Xu, Hao He, Guang-He Lee, Yuyang Wang, and Hao Wang. Graph-relational domain adaptation. *arXiv preprint arXiv:2202.03628*, 2022.

Liang Yang, Xin Chen, Jiaming Zhuo, Di Jin, Chuan Wang, Xiaochun Cao, Zhen Wang, and Yuanfang Guo. Disentangled graph spectral domain adaptation. In *Forty-second International Conference on Machine Learning*, 2025.

Nan Yin, Mengzhu Wang, Zhenghan Chen, Li Shen, Huan Xiong, Bin Gu, and Xiao Luo. Dream: Dual structured exploration with mixup for open-set graph domain adaption. In *International conference on learning representations*, volume 2024, pp. 19776–19797, 2024.

Nan Yin, Li Shen, Mengzhu Wang, Xinwang Liu, Chong Chen, and Xian-Sheng Hua. Dream: a dual variational framework for unsupervised graph domain adaptation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2025.

Yuning You, Tianlong Chen, Zhangyang Wang, and Yang Shen. Graph domain adaptation via theory-grounded spectral regularization. In ... *International Conference on Learning Representations*, volume 2023, pp. https–dblp, 2023.

Haodong Zhang, Tao Ren, Changhu Wang, Yifan Wang, Wei Ju, Huaizhi Tang, Junyu Luo, Zimo Wang, Ziyue Qiao, Xian-Sheng Hua, et al. Disco: Diffusion-guided unbiased discriminative learning for unsupervised graph domain adaptation. In *Proceedings of the 32nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 1*, pp. 1892–1903, 2026.

Xiaowen Zhang, Yuntao Du, Rongbiao Xie, and Chongjun Wang. Adversarial separation network for cross-network node classification. In *Proceedings of the 30th ACM international conference on information & knowledge management*, pp. 2618–2626, 2021.

Zhen Zhang and Bingsheng He. Aggregate to adapt: Node-centric aggregation for multi-source-free graph domain adaptation. In *Proceedings of the ACM on Web Conference 2025*, pp. 4420–4431, 2025.

Zhen Zhang and Bingsheng He. Towards unsupervised open-set graph domain adaptation via dual reprogramming. *Advances in Neural Information Processing Systems*, 38:53320–53343, 2026.

Zhen Zhang, Meihan Liu, Anhui Wang, Hongyang Chen, Zhao Li, Jiajun Bu, and Bingsheng He. Collaborate to adapt: Source-free graph domain adaptation via bi-directional adaptation. In *Proceedings of the ACM Web Conference 2024*, pp. 664–675, 2024.

Jiaming Zhuo, Ziyi Ma, Yintong Lu, Yuwei Liu, Kun Fu, Di Jin, Chuan Wang, Wu Wenning, Zhen Wang, Xiaochun Cao, et al. A closer look at graph transformers: Cross-aggregation and beyond. *Advances in Neural Information Processing Systems*, 38:136781–136807, 2026.

Table 1: Additional structural statistics of the experimental datasets. The statistics are computed using binarized, symmetrized, and self-loop-free graphs.

Dataset	View 1 Avg. Degree	View 2 Avg. Degree	View 1 Homophily	View 2 Homophily	Edge Jaccard
ACM1	8.68	729.83	0.8208	0.6394	0.0069
ACM2	13.39	1078.43	0.8085	0.6393	0.0065
CN	5.70	3.42	0.5057	0.5895	0.0590
US	21.70	10.40	0.5856	0.6485	0.0475
JP	3.00	3.27	0.5867	0.6493	0.0588
FR	4.06	5.04	0.7517	0.6576	0.0768
DE	4.03	5.05	0.5318	0.6185	0.0558
RU	0.73	1.91	0.6746	0.6771	0.0669

A Additional Dataset Statistics

To provide a more detailed characterization of the datasets, we report the average degree, edge homophily, and edge-set Jaccard similarity of different relation views in Table 1. Following the multi-view graph notation introduced in the main text, $\mathcal{E}_v$ denotes the edge set of the v-th relation view. For the statistics reported here, each adjacency matrix A_v is first binarized and symmetrized, after which duplicate edges and self-loops are removed. Thus, $\mathcal{E}_v$ is treated as a set of unique undirected edges.

Based on this edge-set representation, the average degree and edge homophily of the v-th relation view are respectively defined as

$$\bar{d}_v = \frac{2|\mathcal{E}_v|}{|\mathcal{V}|}, \qquad h_v = \frac{1}{|\mathcal{E}_v|} \sum_{\{i,j\}\in\mathcal{E}_v} \mathbb{I}(y_i = y_j), \tag{1}$$

where y_i denotes the class label of node i, and $\mathbb{I}(\cdot)$ is the indicator function. The edge-set Jaccard similarity between two relation views is calculated as

$$J(\mathcal{E}_1, \mathcal{E}_2) = \frac{|\mathcal{E}_1 \cap \mathcal{E}_2|}{|\mathcal{E}_1 \cup \mathcal{E}_2|}. \tag{2}$$

A larger Jaccard similarity indicates greater overlap between the connectivity patterns captured by the two relation views.

In Table 1, Views 1 and 2 correspond to PAP and PSP for ACM1 and ACM2, and to PAP and PP for the remaining MAG datasets, respectively. The labels used to compute homophily are employed only for offline dataset characterization and are never accessed during model training or model selection. Moreover, the edge counts reported in the main text follow the nonzero-entry convention of the original adjacency matrices, whereas the statistics in this table are computed using standardized undirected edge sets. The results reveal substantial structural variation across datasets and relation views. In particular, the PSP views of ACM1 and ACM2 are considerably denser than their corresponding PAP views, whereas the two views of the MAG datasets generally have more comparable average degrees. All edge-set Jaccard similarities are below 0.08, indicating limited edge overlap between the two relation views and suggesting that they capture distinct and potentially complementary connectivity patterns. The differences in homophily further show that the tendency of connected nodes to share the same label varies across both datasets and relation types.

B Model Efficient Experiment

Model Complexity. Here, we analyze the computational complexity of the proposed CARE. Let N denote the total number of nodes in the source domain and target domain, V the number of relation views, E the total number of observed edges, d the input feature dimension, h the hidden dimension, p the embedding dimension, and K the number of APPNP propagation steps. Edge reweighting requires $\mathcal{O}(VNdh + Eh)$ operations, while K-step sparse APPNP propagation requires $\mathcal{O}(KEd)$. Cross-domain alignment over corresponding-view and cross-view pairs requires $\mathcal{O}(V^2 B_S B_T p)$, where B_S and B_T denote the source and target batch sizes. Combining these terms yields an overall per-iteration complexity of $\mathcal{O}(VNdh+E(h+Kd)+V^2 B_S B_T p)$. Under full-batch

alignment, the dominant term becomes $\mathcal{O}(V^2N^2p)$. Since the number of relation views is typically small, CARE remains comparable to existing MGDA methods in asymptotic complexity and does not introduce a higher order of complexity with respect to the number of nodes.

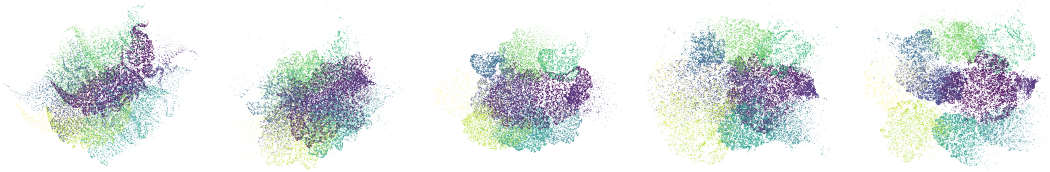
Model Efficient Experiment. To further evaluate the computational efficiency of CARE, we compare its training time and GPU memory usage with representative GDA and MGDA baselines, including UDAGCN, ASN, PA, HGDA, and SAGA. Following SAGA, both training time and memory usage are normalized with respect to CARE on each transfer task. We also report the classification accuracy to jointly evaluate model efficiency and predictive performance. The results are summarized in Table 2.

Table 2: Comparison of Training Time, Memory Usage, and Accuracy on ACM and MAG datasets.

Dataset	Method	Training Time (Normalized w.r.t. CARE)	Memory Usage (Normalized w.r.t. CARE)	Accuracy (%)
ACM1 → ACM2	CARE	1.000	1.000	53.8
	UDAGCN	0.784	0.760	45.2
	ASN	1.030	1.070	43.6
	PA	0.928	0.867	50.6
	HGDA	0.716	0.827	51.1
	SAGA	0.895	0.860	52.3
CN → US	CARE	1.000	1.000	59.3
	UDAGCN	0.862	0.821	42.5
	ASN	1.218	1.307	41.6
	PA	0.845	0.952	40.2
	HGDA	0.858	1.031	41.7
	SAGA	0.947	0.965	55.3
JP → CN	CARE	1.000	1.000	53.2
	UDAGCN	0.928	0.964	45.1
	ASN	1.308	1.467	33.5
	PA	0.924	1.072	35.1
	HGDA	0.933	1.160	39.2
	SAGA	1.035	1.066	51.2
DE → FR	CARE	1.000	1.000	54.3
	UDAGCN	0.890	0.897	35.2
	ASN	1.188	1.418	43.3
	PA	0.895	1.035	41.6
	HGDA	0.902	1.045	39.1
	SAGA	1.009	1.044	52.9
RU → US	CARE	1.000	1.000	38.9
	UDAGCN	0.830	0.869	35.9
	ASN	1.063	1.275	28.9
	PA	0.835	0.963	30.6
	HGDA	0.847	0.935	36.1
	SAGA	0.928	0.953	37.7
CN → DE	CARE	1.000	1.000	38.9
	UDAGCN	0.876	0.949	36.9
	ASN	1.151	1.435	37.2
	PA	0.936	1.099	35.7
	HGDA	0.935	1.173	31.7
	SAGA	1.036	1.095	38.1

C Experimental Setup

The experiments are implemented in PyTorch 2.8.0 on a server equipped with an Intel(R) Xeon(R) Gold 6248R CPU @ 3.00GHz and an NVIDIA GeForce RTX 3090 24GB GPU. For CARE, we first perform graph propagation using APPNP and then employ a two-layer MLP encoder. The hidden and embedding dimensions are set to 256 and 128, respectively, and the dropout rate is set to 0.3. The model and the edge reweighting module are jointly optimized using Adam (Kingma & Ba, 2014), with the learning rates of both modules set to 1×10^{-4}. The weight decay is selected from $\{1 \times 10^{-4}, \ldots, 5 \times 10^{-3}\}$. For fairness, we use the same parameter settings for all cross-domain node classification methods in our experiments, except for some special cases. We repeat each experiment with five random seeds and report the average ACC and Macro-F1 scores.



(a) UDAGCN (b) ASN (c) PA (d) SAGA (e) CARE

Figure 1: Visualizing source and target node embeddings via t-SNE. Each color represents a class, while light and dark shades of the same color correspond to nodes of the same class from the source domain and target domain, respectively.

D VISUALIZATION

To evaluate the quality of learned representations, we visualize the node embeddings generated by CARE and compare them with UDAGCN, PA, ASN, and SAGA using t-SNE on the MAG datasets. As shown in Figure 1, each color represents a different node class, while the lighter and darker shades within each class denote nodes from the source domain and target domain, respectively. Compared with the baseline methods, CARE produces more clearly separated clusters for different classes, while source and target nodes within the same class show greater overlap. This indicates that CARE learns representations with better class discrimination and cross-domain alignment.

E DESCRIPTION OF ALGORITHM CARE

The overall optimization procedure of CARE is summarized in Algorithm 1.

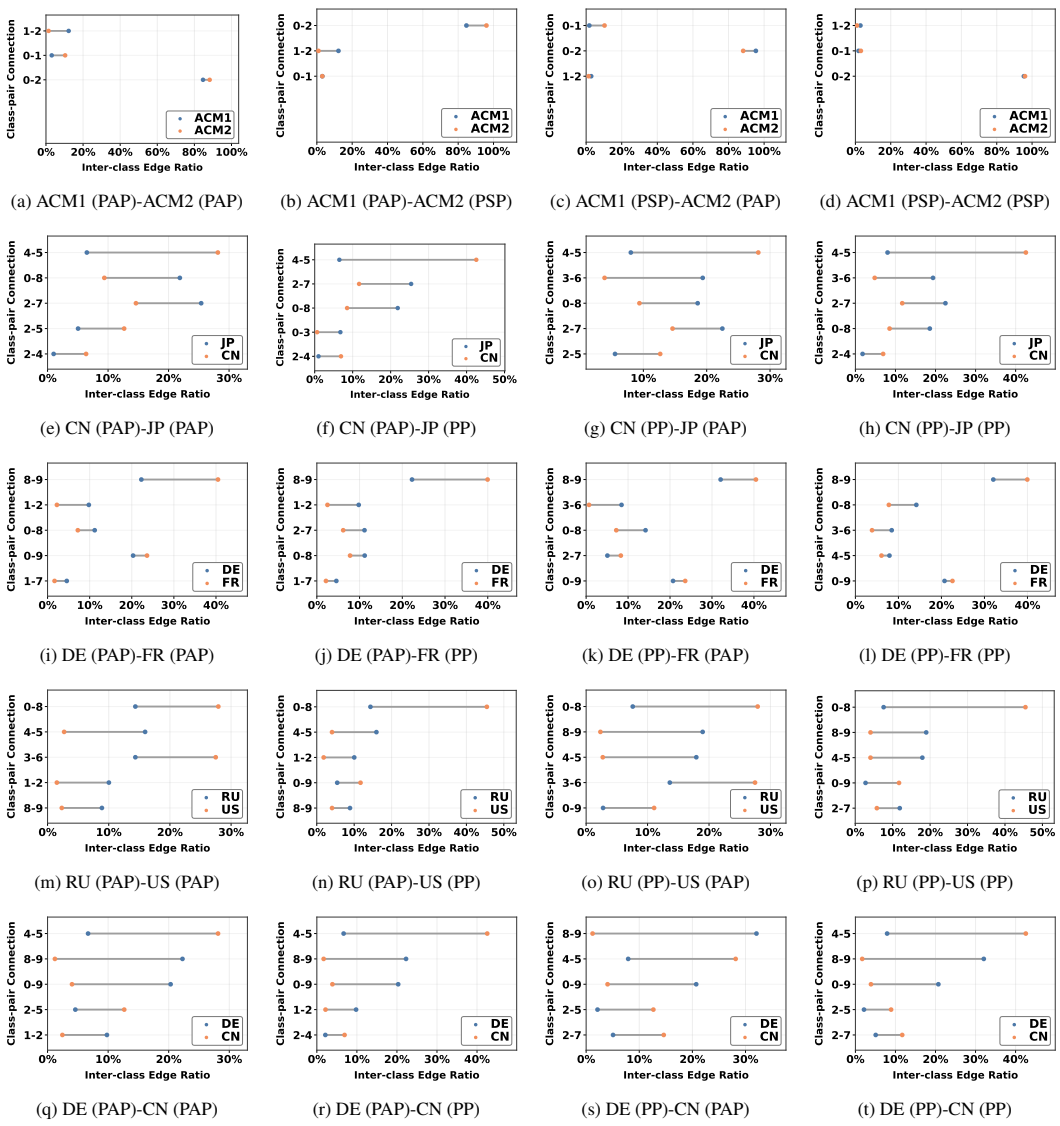
Algorithm 1 The proposed algorithm CARE

Require: Source multi-view graph G^S; target multi-view graph G^T; source node label matrix Y^S; maximum number of iterations η
Ensure: Classification result $\hat{Y}^T$

for $it = 1$ to η **do**
  for $D \in \{S, T\}$ and $v = 1, \ldots, V$ **do**
    Compute edge scores and reweighting factors for A_v^D
    Construct the reweighted graph $\tilde{G}_v^D$
    Normalize $\tilde{A}_v^D$ and perform K-step APPNP propagation
    Encode $\tilde{X}_v^{D,(K)}$ to obtain $\tilde{Z}_v^D$ using the shared encoder
  end for
  Project source and target representations into the alignment space
  Compute bidirectional alignment for corresponding-view and cross-view pairs
  Aggregate the alignment losses to obtain $\mathcal{L}_{\mathrm{A}}$
  Obtain source and target predictions from the multi-view representations
  Compute the source classification loss $\mathcal{L}_S$
  Compute the target entropy loss $\mathcal{L}_T$
  Calculate the overall loss $\mathcal{L}$ according to Eq. (14)
  Update all parameters of CARE according to $\mathcal{L}$
end for
Predict the labels of target graph nodes using the trained model
return $\hat{Y}^T$

F EMPIRICAL ANALYSIS OF CROSS-DOMAIN EDGE DISTRIBUTION SHIFTS

To further investigate cross-domain edge distribution shifts across different relation views, we extend the analysis in the main text to additional ACM and MAG domain pairs. For each pair of



(a) ACM1 (PAP)-ACM2 (PAP) (b) ACM1 (PAP)-ACM2 (PSP) (c) ACM1 (PSP)-ACM2 (PAP) (d) ACM1 (PSP)-ACM2 (PSP)

(e) CN (PAP)-JP (PAP) (f) CN (PAP)-JP (PP) (g) CN (PP)-JP (PAP) (h) CN (PP)-JP (PP)

(i) DE (PAP)-FR (PAP) (j) DE (PAP)-FR (PP) (k) DE (PP)-FR (PAP) (l) DE (PP)-FR (PP)

(m) RU (PAP)-US (PAP) (n) RU (PAP)-US (PP) (o) RU (PP)-US (PAP) (p) RU (PP)-US (PP)

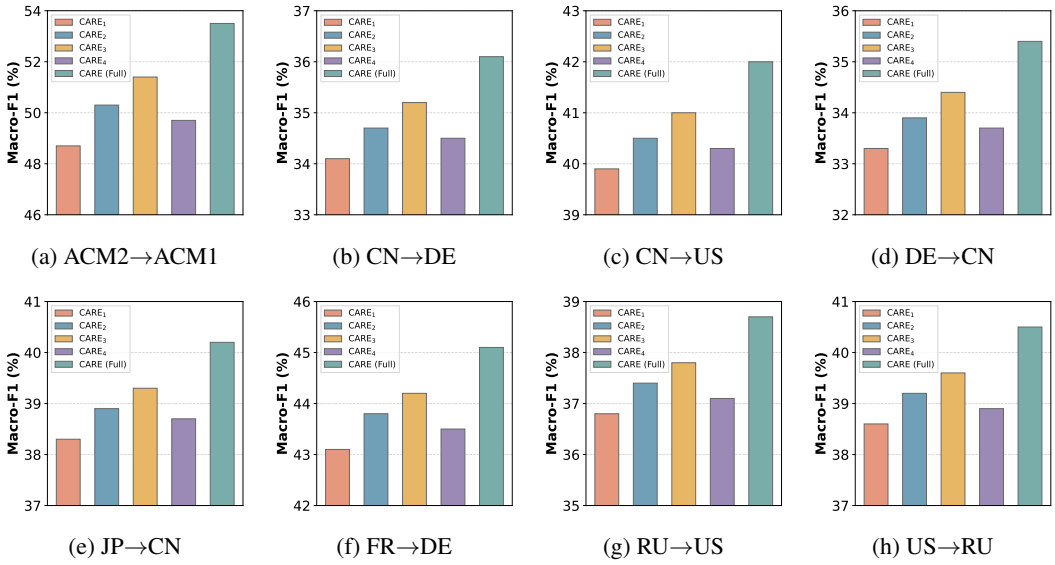
(q) DE (PAP)-CN (PAP) (r) DE (PAP)-CN (PP) (s) DE (PP)-CN (PAP) (t) DE (PP)-CN (PP)

Figure 2: Cross-Domain Discrepancies in Class-Pair Edge Distributions across Different Domains and Relation Views.

source and target views, the edge ratio represents the fraction of inter-class edges connecting each class pair, excluding intra-class edges. We measure the cross-domain discrepancy by the absolute difference between the corresponding edge ratios in the source and target views, with longer lines indicating larger discrepancies. As shown in Figure 2, different view combinations exhibit different discrepancy patterns even for the same pair of domains. These results further show that different relation views exhibit distinct cross-domain edge distribution shifts.

G Ablation Study Analysis

We provide additional ablation results to further examine the contribution of different components in CARE. In addition to the representative transfer tasks reported in the main text, Figure 3 presents the results on the remaining transfer tasks. Overall, removing individual components consistently affects the performance of CARE, while the magnitude of performance degradation varies across datasets.



(a) ACM2→ACM1 (b) CN→DE (c) CN→US (d) DE→CN

(e) JP→CN (f) FR→DE (g) RU→US (h) US→RU

Figure 3: Additional ablation results of CARE on different transfer tasks.

ACM Datasets. The ablation results on ACM can be understood from its structural characteristics. ACM contains only three classes and therefore has only three possible inter-class edge types, with a large fraction of inter-class edges concentrated in a small number of class pairs. At the same time, PAP and PSP show clear differences in connectivity: their Edge Jaccard similarity is below 1%, while PSP has a much larger average degree than PAP. Under this setting, edge reweighting and cross-domain alignment are important for controlling how information is propagated across the two relation views. Removing any of these components may weaken the model's ability to preserve useful cross-domain information during message propagation.

MAG Datasets. MAG presents a different structural pattern. With ten classes, its inter-class edges are distributed over a much larger number of class pairs, and the class-pair edge distribution is more dispersed than in ACM. In addition, the two relation views generally have more comparable average degrees and higher Edge Jaccard similarity. This provides multiple sources of structural information across views and class-pair connections. As a result, different components of CARE can complement each other during cross-domain adaptation, while the complete model benefits from jointly combining edge reweighting, cross-domain guidance, cross-view alignment, and source-target reweighting.

H Parameter Analysis

CARE mainly involves five hyperparameters: λ, ρ, γ, δ, and α. The parameter λ controls the contribution of cross-view alignment, while ρ determines the maximum magnitude of edge reweighting. The parameters γ and δ balance the cross-domain alignment loss and the target-domain entropy loss, respectively. The parameter α is the teleport probability in APPNP and controls the amount of initial node features retained during propagation. In our experiments, λ is selected from the set $\{0, 0.05, 0.1, 0.5, 1.0\}$, ρ from the set $\{0.1, 0.2, 0.3, 0.4, 0.5\}$, γ from the set $\{0.01, 0.03, 0.05, 0.08, 0.1\}$, δ from the set $\{0.001, 0.01, 0.05, 0.1, 1.0\}$, and α from the set $\{0, 0.03, 0.05, 0.1, 0.5\}$. The sensitivity of CARE to these hyperparameters is shown in Figure 4.

ACM Datasets. The two relation views in ACM exhibit pronounced structural differences. As shown in Table 1, PAP is relatively sparse, whereas PSP has a substantially larger average degree, and the Edge Jaccard similarity between the two views is below 1%. Meanwhile, both views exhibit relatively high homophily. Considering the large inter-view structural differences, λ and ρ are searched within relatively small ranges, namely $\lambda \in \{0, 0.05, 0.1\}$ and $\rho \in \{0.1, 0.2, 0.3\}$, to avoid overly strong cross-view alignment and excessive edge-weight adjustment. Since the relatively high homophily provides reliable neighborhood information for propagation, α is also searched

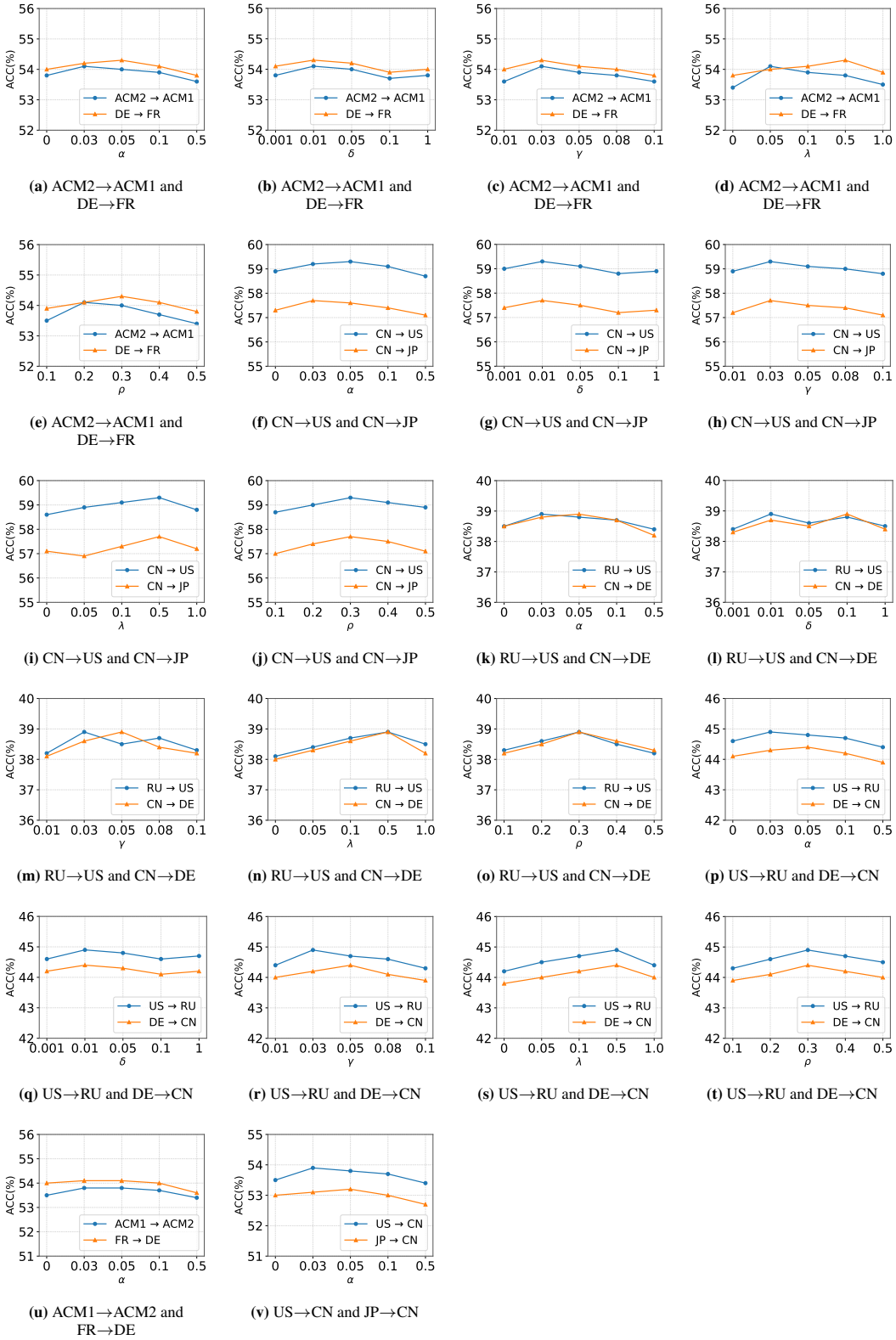
within a smaller range of $\{0, 0.03, 0.05\}$, allowing APPNP to place more emphasis on neighborhood aggregation.

MAG Datasets. Compared with ACM, the two relation views in MAG generally have more comparable average degrees and higher Edge Jaccard similarity, while their graph sparsity and homophily vary considerably across different domains. Therefore, stronger cross-view interaction and a larger edge-reweighting range can be considered, and λ and ρ are searched within relatively larger ranges of $\{0.1, 0.5, 1.0\}$ and $\{0.2, 0.3, 0.4, 0.5\}$, respectively. In addition, the MAG graphs are generally sparser and exhibit more variable homophily than ACM, making it useful to retain more initial node information during propagation. Accordingly, α is searched within the larger-value subset $\{0.03, 0.05, 0.1, 0.5\}$. The parameters γ and δ are selected from their full candidate sets for both ACM and MAG to balance cross-domain alignment and target-domain entropy regularization across different transfer tasks.

Table 3: Hyperparameter settings for different transfer tasks.

Types	Datasets	λ	ρ	γ	δ	α
ACM	ACM1→ACM2	0.1	0.2	0.03	0.01	0.03
	ACM2→ACM1	0.1	0.2	0.03	0.01	0.03
MAG	CN→US	0.5	0.3	0.03	0.01	0.05
	US→CN	0.5	0.3	0.03	0.01	0.03
	JP→CN	0.5	0.2	0.03	0.05	0.05
	CN→JP	0.5	0.3	0.03	0.01	0.03
	DE→FR	0.5	0.3	0.03	0.01	0.05
	FR→DE	0.1	0.3	0.03	0.01	0.03
	RU→US	0.5	0.3	0.03	0.01	0.03
	US→RU	0.5	0.3	0.03	0.01	0.03
	DE→CN	0.5	0.3	0.05	0.01	0.05
	CN→DE	0.5	0.3	0.05	0.1	0.05

Overall, CARE maintains relatively stable performance across a broad range of hyperparameter values, as shown in Figure 4. The optimal hyperparameter settings adopted for individual transfer tasks are summarized in Table 3.



(a) ACM2→ACM1 and DE→FR **(b)** ACM2→ACM1 and DE→FR **(c)** ACM2→ACM1 and DE→FR **(d)** ACM2→ACM1 and DE→FR

(e) ACM2→ACM1 and DE→FR **(f)** CN→US and CN→JP **(g)** CN→US and CN→JP **(h)** CN→US and CN→JP

(i) CN→US and CN→JP **(j)** CN→US and CN→JP **(k)** RU→US and CN→DE **(l)** RU→US and CN→DE

(m) RU→US and CN→DE **(n)** RU→US and CN→DE **(o)** RU→US and CN→DE **(p)** US→RU and DE→CN

(q) US→RU and DE→CN **(r)** US→RU and DE→CN **(s)** US→RU and DE→CN **(t)** US→RU and DE→CN

(u) ACM1→ACM2 and FR→DE **(v)** US→CN and JP→CN

Figure 4: Hyperparameter sensitivity analysis of CARE on the remaining transfer tasks.