FROM EUCLIDEAN TO RIEMANNIAN: BOUNDARY-AWARE GEODESIC OVERSAMPLING FOR CLASS-IMBALANCED GRAPHS

Anonymous authors
Paper under double-blind review

ABSTRACT

Class imbalance is a widespread problem in graph-structured data. A promising solution is to oversample minority samples to balance the class distribution. However, existing methods mainly synthesize samples via Euclidean linear interpolation, which deviates from the intrinsic geometry of graph representations. Moreover, even the interpolation between intra-category samples passes through regions dominated by other categories, while improperly synthesized edges further distort the local topology. To address these limitations, we propose a **B**oundary-**A**ware **R**iemannian **G**eodesic **O**versampling method (**BARGO**) for class-imbalanced graph learning. Specifically, we first employ a Riemannian graph encoder to map nodes onto a geometry tailored to the hierarchical structures of graphs and construct Fréchet prototypes to characterize class-wise distributions. Then, we sample reliable minority seed and partner nodes, and synthesize candidate nodes along their intrinsic geodesics. Furthermore, we develop a dual screening method that evaluates both endpoint margins and intermediate points along the generation path, thereby reducing the risk of crossing category boundaries. Finally, we propose a geometry-aware topology generator. Experiments on six real-world datasets demonstrate the effectiveness of the proposed method.

1 INTRODUCTION

Graph Neural Networks (GNNs) have achieved impressive success in graph representation learning (Wu et al., 2020). They have been widely applied in various graph analyzing tasks, including node classification (Yang et al., 2024), link prediction (Zhang & Chen, 2018), graph classification (Xie et al., 2022), etc. Among these tasks, node classification has attracted particular attention in fraud detection (Li et al., 2022), transportation analysis (Rahmani et al., 2023), fault diagnosis (Chen et al., 2023a), and so on.

However, the success of GNNs commonly relies on a balanced distribution of labeled nodes across various categories (Ju et al., 2025). This assumption rarely holds in real-world scenarios. For example, in the task of fake user detection, the vast users are normal, while fake ones represent only a tiny fraction. Once trapped in the problem of class imbalance, the majority classes are over-fitting, and the minority ones are under-represented (Li et al., 2024).

Existing methods for class-imbalanced graph learning can be broadly categorized into loss modifying methods (Wang et al., 2024; Chen et al., 2023b; Peng et al., 2023) and oversampling methods (Wang et al., 2020; 2022; Xie et al., 2021). The former aims to assign greater importance to minority samples during optimization. This paradigm is simple and effective, but is constrained by the limited number of labels and cannot alleviate the issue of insufficient supervision signal for minority samples. The latter instead intends to synthesize samples to balance the quantity of samples in each category. This paradigm effectively enriches the supervision signal by increasing the number of minority samples and avoids overfitting issues. For example, GraphSMOTE interpolates minority representations and predicts edges for synthesized nodes (Zhao et al., 2021). GraphENS synthesizes ego networks for minority classes with a saliency-based node mixing technique, emphasizing the message-passing mechanism (Park et al., 2021). GraphSHA synthesizes harder minority samples to enlarge the decision boundaries of minority classes (Li et al., 2023).

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.

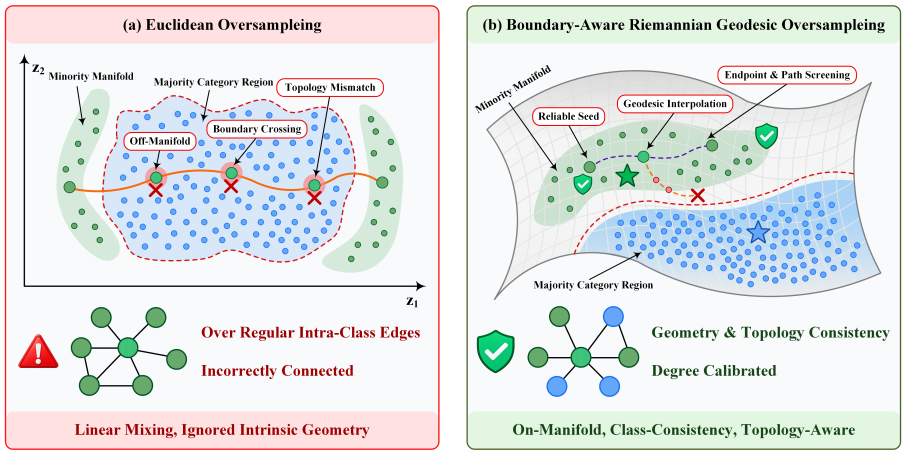


Figure 1: Illustration of Euclidean and Riemannian oversampling. Euclidean linear interpolation leaves the intrinsic minority-class manifold, cross majority-dominated regions, and introduces inconsistent topology. In contrast, our method performs geodesic interpolation on the Riemannian manifold and employs endpoint- and path-level screening to synthesize geometrically valid, class-consistent, and topology-aware minority nodes.

Despite their effectiveness, most existing oversampling methods on graphs perform linear interpolation in either the input space or a Euclidean latent space. This design overlooks the fact that real-world graphs typically exhibit non-Euclidean geometry. Riemannian space enables rapidly expanding capacity and has demonstrated advantages in representing non-Euclidean structures with limited distortion (Chami et al., 2019; Liu et al., 2019). Direct Euclidean interpolation deviates from the underlying data manifold and synthesizes geometrically unrealistic samples, as illustrated in Figure 1 (a). Furthermore, replacing Euclidean interpolation with geodesic interpolation alone cannot guarantee the reliability of oversampling. We emphasize that oversampling on Riemannian space encounters the following two challenges.

(1) The risk of cross-boundary geodesic interpolation. Existing interpolation-based methods assume that a path linking two minority samples remains inside the corresponding category region. However, category regions learned from imbalanced graphs are non-convex and distorted by majority dominated message-passing. Even both endpoints belong to the same class, their linking geodesic may traverse an ambiguous or majority dominated region. Once noisy or boundary-adjacent samples serve as interpolation endpoints, this issue is further amplified. Therefore, reliable oversampling on a Riemannian space requires geometrically valid interpolation and explicit control over seed reliability, endpoint margins, and path generation.

(2) The inconsistency between synthetic representations and topology. Unlike independent samples in Euclidean data, a synthesized node is not an isolated feature vector. Its representation is coupled with the topology. However, existing topology modeling methods typically connect the synthesized node to nearby or intra-class nodes. This leads to overly regular neighborhoods, whereas unconstrained edge prediction introduces inconsistent structural roles. Thus, oversampling on a Riemannian space necessitates topology modeling that coordinates node synthesis with neighborhood inheritance, degree distribution, and structural similarity.

To this end, we propose a **B**oundary-**A**ware **R**iemannian **G**eodesic **O**versampling method (**BARGO**) for class-imbalanced graph learning, as illustrated in Figure 1 (b). Moving beyond Euclidean linear interpolation, BARGO encodes samples on a Riemannian manifold and characterizes each category with a Fréchet prototype. It samples reliable minority seeds and partners, and then synthesizes candidates along shortest geodesics via exponential and logarithmic maps. To avoid the issue of cross-boundary synthesis, we devise a two-level screening method that evaluates both candidate endpoints and intermediate path points. It ensures that unreliable candidates are excluded, redirected toward the target prototype. For reliable ones, we design a geometry-aware topology generator that evaluates inherited candidate neighbors for topology modeling. We conduct experiments on six

real-world datasets. The experimental results demonstrate that the proposed method outperforms the competitive methods for class-imbalanced issues.

The main contributions are summarized as follows:

- We propose a boundary-aware geodesic oversampling method that couples reliable endpoint sampling with dual screening, reducing the risk of synthesizing minority samples across ambiguous category regions.
- We develop a geometry-aware coordinated topology modeling method that samples neighborhoods for synthesized nodes while maintaining degree and structural consistency.
- We conduct experiments on six real-world datasets. The experimental results demonstrate that the proposed method consistently improves graph representation learning under class-imbalanced scenarios.

2 Related Work

2.1 Class-Imbalanced Learning on Graphs

The impressive performance of GNNs typically relies on a balanced class distribution (Liu et al., 2025b). However, real-world graphs often exhibit imbalanced distributions. Existing methods can be broadly categorized into loss modifying and oversampling methods. The former aims to focus on the minority samples by distributing larger training weights. For instance, Focal Loss reduces the focus of the model on easily classified samples and concentrates more on difficult minority samples (Lin et al., 2017). CB Loss weights classes according to their effective sample numbers (Cui et al., 2019). PC Softmax corrects prediction logits by class priors without modifying model training (Hong et al., 2021). RDR dynamically estimates class density ratios to adapt weights to evolving feature distributions (Luo et al., 2024). The latter explicitly adds the number of minority samples to tackle the problem of class imbalance. For example, GraphSMOTE interpolates minority embeddings and reconstructs edges through a link predictor (Zhao et al., 2021). GraphENS synthesizes minority ego networks to alleviate the issues of neighbor memorization (Park et al., 2021). GraphSHA synthesizes hard minority samples near decision boundaries to mitigate the compression of minority subspace (Li et al., 2023). CDCGAN further models class distributions with a conditional generative adversarial paradigm (Liu et al., 2025a). Furthermore, researchers also explored contrastive-based approaches to alleviate the class imbalance on graphs (Zeng et al., 2023).

2.2 Riemannian Graph Learning

Graph-structured data often exhibit hierarchical patterns that cannot be represented in Euclidean spaces without distortion (Peng et al., 2021). Riemannian graph learning alleviates this limitation by embedding graph structures into manifolds with geometric inductive biases (Fei et al., 2025). For example, HGCN maps Euclidean node features into hyperbolic spaces and performs feature transformation and neighborhood aggregation in tangent spaces with trainable layer-wise curvature (Chami et al., 2019). HGNN develops a manifold-agnostic method and explores graph representation learning in both the Poincaré and Lorentz models (Liu et al., 2019). To preserve hyperbolic geometry throughout message-passing, LGCN introduces Lorentzian feature transformations and neighborhood aggregation on the basis of the Lorentzian centroid (Zhang et al., 2021). Beyond hyperbolic spaces, κ-GCN provides a unified formulation of graph convolution in various spaces, with curvature learned according to the underlying graph structure (Bachmann et al., 2020). SelfMGNN further combines multiple Riemannian component spaces through a Cartesian product and employs hierarchical attention to capture heterogeneous structural patterns (Sun et al., 2022).

3 Methodology

3.1 Preliminary

Definition 1. **Attributed Graph.** Let $\mathcal{G} = (\boldsymbol{\mathcal{V}}, \boldsymbol{\mathcal{E}}, \boldsymbol{X})$ denote an attributed graph, where $\boldsymbol{\mathcal{V}} = \{v_1, \ldots, v_n\}$ is the set of nodes, $\boldsymbol{\mathcal{E}} = \{e_1, \ldots, e_m\}$ is the set of edges, and $\boldsymbol{X} \in \mathbb{R}^{n \times d}$ is the feature

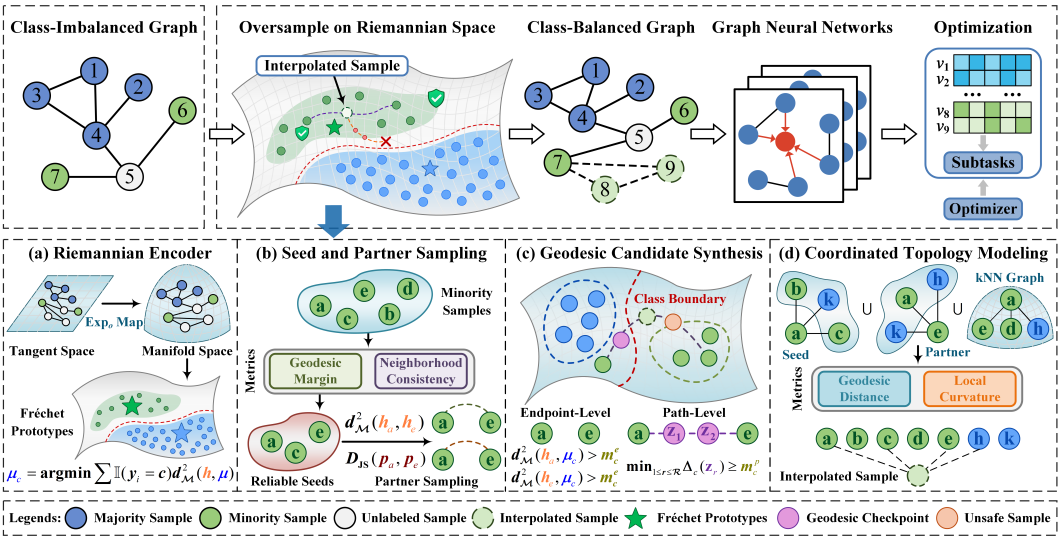


Figure 2: The framework of the proposed BARGO.

matrix. The adjacency matrix is denoted by $\boldsymbol{A} \in \{0,1\}^{n \times n}$. The nodes are divided into a labeled set $\boldsymbol{\mathcal{V}}^L$ and an unlabeled set $\boldsymbol{\mathcal{V}}^U$, where $|\boldsymbol{\mathcal{V}}^L| \ll |\boldsymbol{\mathcal{V}}^U|$. Each labeled node $v_i \in \boldsymbol{\mathcal{V}}^L$ is associated with a label $y_i \in \boldsymbol{C} = \{1, \ldots, c\}$.

Definition 2. **Class Imbalance.** Let $n_c = |\{v_i \in \boldsymbol{\mathcal{V}}^L | y_i = c\}|$ be the number of labeled nodes in class c. The imbalance ratio ρ is defined as

$$\rho = \frac{\max_{n_c}}{\min_{n_c}}. \tag{1}$$

A larger ρ indicates a more severely imbalanced label distribution.

Definition 3. **Riemannian Manifold.** Let $\mathcal{M}$ be a Riemannian manifold equipped with a metric tensor and the induced geodesic distance $d_{\mathcal{M}}(\cdot, \cdot)$. The exponential map $\text{Exp}_p : T_p\mathcal{M} \to \mathcal{M}$ maps a tangent vector at p onto the manifold, while the logarithmic map $\text{Log}_p : \mathcal{M} \to T_p\mathcal{M}$ performs the inverse operation locally. We primarily instantiate $\mathcal{M}$ as a Lorentz manifold, while the proposed framework is applicable to other differentiable manifolds.

Problem Definition. Given an imbalanced attributed graph $\mathcal{G}$, the goal is to generate geometrically valid and topologically compatible minority nodes, construct a class-balanced graph $\widetilde{\mathcal{G}}$, and learn a classifier f_θ that predicts the labels of nodes in $\boldsymbol{\mathcal{V}}^U$.

3.2 Framework

We propose **BARGO**, a **B**oundary-**A**ware **R**iemannian **G**eodesic **O**versampling method, as illustrated in Figure 2. It consists of four main modules: (1) Riemannian graph representation learning, which maps nodes onto a manifold compatible with the intrinsic graph geometry; (2) reliable seed and partner sampling, which prevents noisy and boundary minority nodes from being repeatedly amplified; (3) boundary-aware geodesic synthesis, which synthesizes candidates along intrinsic geodesics and evaluates both their endpoints and interpolation paths; and (4) geometry-aware coordinated topology modeling, which establishes reliable links for synthetic nodes.

3.3 Riemannian Graph Representation Learning

Riemannian Graph Encoder. The node features in the Euclidean space are first transformed into the tangent space at a reference point $o \in \mathcal{M}$ and then mapped onto the manifold:

$$\boldsymbol{h}_i = \text{Exp}_o\left(\boldsymbol{W}_0 \boldsymbol{x_i} + \boldsymbol{b}_0\right), \tag{2}$$

where $\boldsymbol{x_i}$ is the feature of node v_i, $\boldsymbol{W}_0$ and $\boldsymbol{b}_0$ are learnable parameters, $\boldsymbol{h}_i$ is the manifold representation of v_i.

At the l-th layer, BARGO aggregates neighboring representations in a common tangent space:

$$\boldsymbol{m}_i^{(l)} = \sum_{j \in \mathcal{N}(i) \cup i} \alpha_{ij}^{(l)} \boldsymbol{W}^{(l)} \operatorname{Log}_o \left(\boldsymbol{h}_j^{(l)} \right), \tag{3}$$

$$\boldsymbol{h}_i^{(l+1)} = \operatorname{Exp}_o \left[\phi \left(\boldsymbol{m}_i^{(l)} \right) + \operatorname{Log}_o \left(\boldsymbol{h}_i^{(l)} \right) \right], \tag{4}$$

where $\mathcal{N}(i)$ denotes the neighborhood of v_i, $\alpha_{ij}^{(l)}$ is a normalized aggregation coefficient, $\boldsymbol{m}_i^{(l)}$ is the neighborhood information, and $\phi(\cdot)$ is an activation function. The residual term alleviates over-smoothing and reduces the risk that minority representations are dominated by majority neighbors.

Fréchet Category Prototypes. For each category c, we summarize its labeled representations using a Fréchet prototype:

$$\boldsymbol{\mu}_c = \arg\min_{\boldsymbol{\mu} \in \mathcal{M}} \sum_{i \in \mathcal{V}_L} \mathbb{I}(y_i = c) d_{\mathcal{M}}^2(\boldsymbol{h}_i, \boldsymbol{\mu}). \tag{5}$$

The probability that node v_i belongs to class c is evaluated using its geodesic distance to the category prototypes:

$$p(y_i = c \mid \boldsymbol{h}_i) = \frac{\exp\left[-d_{\mathcal{M}}^2(\boldsymbol{h}_i, \boldsymbol{\mu}_c)/\tau\right]}{\sum_{k=1}^{C} \exp\left[-d_{\mathcal{M}}^2(\boldsymbol{h}_i, \boldsymbol{\mu}_k)/\tau\right]}, \tag{6}$$

where $\tau > 0$ is a temperature parameter. Compared to the Euclidean-based category prototype, the Fréchet prototype respects the geometry of $\mathcal{M}$ and provides a natural reference for evaluating class consistency.

3.4 Reliable Seed and Partner Sampling

The randomly sampled minority nodes amplify label noise and yield candidates across ambiguous regions. Therefore, the proposed BARGO evaluates the reliability of each minority node from the perspectives of geometry and topology. Specifically, the geodesic category margin of node v_i is defined as:

$$g_i = \min_{c \neq y_i} d_{\mathcal{M}}(\boldsymbol{h}_i, \boldsymbol{\mu}_c) - d_{\mathcal{M}}(\boldsymbol{h}_i, \boldsymbol{\mu}_{y_i}). \tag{7}$$

where $\boldsymbol{\mu}_c$ denotes the Fréchet prototypes of class c. A positive g_i indicates that the node v_i is closer to its own prototype than to any competing one. We further measure its neighborhood consistency by:

$$q_i = \frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} p_{j,y_i}, \tag{8}$$

where p_{j,y_i} denotes the predicted probability that v_i belongs to class y_i. The overall reliability score is defined as:

$$r_i = p_{i,y_i} \cdot \sigma\left(\frac{g_i}{\tau_g}\right) q_i, \tag{9}$$

where $\sigma(\cdot)$ denotes the sigmoid function and $\tau_g > 0$ controls the sensitivity to the geodesic margin. For each minority class c, the reliable seed set $\mathcal{S}_c$ is constructed as:

$$\boldsymbol{\mathcal{S}}_c = \{v_i \in \boldsymbol{\mathcal{V}}^L : y_i = c, \hat{y}_i = y_i, g_i > 0, r_i \geq \eta_r\} \tag{10}$$

where η_r is the threshold of reliability score.

For each seed node $v_i \in \boldsymbol{\mathcal{S}}_c$, we construct a distance-window partner set:

$$\boldsymbol{\mathcal{P}}_i = \{v_j \in \boldsymbol{\mathcal{S}}_c \setminus v_i : \delta_c^- \leq d_{\mathcal{M}}(\boldsymbol{h}_i, \boldsymbol{h}_j) \leq \delta_c^+\}. \tag{11}$$

where $0 \leq \delta_c^- \leq \delta_c^+$. The lower bound prevents nearly duplicated samples, whereas the upper bound reduces the risk of interpolating between distant modes through ambiguous regions. A partner $v_j \in \boldsymbol{\mathcal{P}}_i$ is sampled according to

$$\pi_{ij} \propto \exp\left(-\frac{d_{\mathcal{M}}^2(\boldsymbol{h}_i, \boldsymbol{h}_j)}{\tau_d} - \frac{D_{\mathrm{JS}}(\boldsymbol{p}_i, \boldsymbol{p}_j)}{\tau_p}\right), \tag{12}$$

where $D_{\mathrm{JS}}(\boldsymbol{p}_i, \boldsymbol{p}_j)$ measures the divergence between their predictive distributions, while $\tau_d > 0$ and $\tau_p > 0$ control the sensitivities to geometric distance and semantic discrepancy, respectively. Seeds with an empty partner set are excluded from the current oversampling round.

3.5 BOUNDARY-AWARE GEODESIC SYNTHESIS

Geodesic Candidate Synthesis. Given a reliable intra-class pair (v_i, v_j), BARGO synthesizes a candidate along their geodesic as:

$$\widetilde{\boldsymbol{h}} = \mathrm{Exp}_{\boldsymbol{h}_i}\left(t \,\mathrm{Log}_{\boldsymbol{h}_i}(\boldsymbol{h}_j)\right), \qquad t \sim \mathrm{Beta}(a, b), \tag{13}$$

where $a, b > 0$. Unlike Euclidean linear interpolation, this operation generates a point on the manifold along the intrinsic geodesic connecting the seed and partner nodes. The target number of nodes for class c is determined as:

$$n_c^* = \min\left\{n_{\max}, \left\lceil n_c^{1-\gamma} n_{\max}^{\gamma} \right\rceil\right\}, \tag{14}$$

where $n_{\max} = \max_{n_c}$ and $\gamma \in [0, 1]$ controls the balancing strength. Accordingly, BARGO repeatedly synthesizes and screens candidates until $S_c = \max\{0, n_c^* - n_c\}$ candidates are accepted for class c.

Endpoint-Level Screening. The intra-class endpoints do not guarantee that the synthesized candidate remains inside the corresponding category region. We therefore define its prototype margin as:

$$\Delta_c(\widetilde{\boldsymbol{h}}) = \min_{k \neq c} d_{\mathcal{M}}\left(\widetilde{\boldsymbol{h}}, \boldsymbol{\mu}_k\right) - d_{\mathcal{M}}\left(\widetilde{\boldsymbol{h}}, \boldsymbol{\mu}_c\right). \tag{15}$$

Let $\widetilde{\boldsymbol{p}} = [\widetilde{p}_1, \ldots, \widetilde{p}_C]$ denote the predictive distribution of the candidate. Its predictive entropy is

$$\mathcal{H}(\widetilde{\boldsymbol{p}}) = -\sum_{k=1}^{C} \widetilde{p}_k \log \widetilde{p}_k. \tag{16}$$

A candidate passes endpoint-level screening if

$$\Delta_c(\widetilde{\boldsymbol{h}}) \geq m_c^e, \qquad \widetilde{p}_c \geq \eta_e, \qquad \mathcal{H}(\widetilde{\boldsymbol{p}}) \leq \xi_e, \tag{17}$$

where m_c^e is a category-specific safety margin, η_e is the minimum probability of the target-category, and ξ_e is the maximum predictive-entropy threshold.

Path-Level Screening. A safe endpoint does not ensure a safe interpolation trajectory. We therefore sample $\mathcal{R}$ intermediate checkpoints along the geodesic:

$$\boldsymbol{z}_r = \mathrm{Exp}_{\boldsymbol{h}_i}\left(s_r t \,\mathrm{Log}_{\boldsymbol{h}_i}(\boldsymbol{h}_j)\right), \qquad s_r = \frac{r}{\mathcal{R}+1}, \quad r = 1, \ldots, \mathcal{R}. \tag{18}$$

Since the seed has passed the reliability screening and the candidate has passed the endpoint-level screening, the intermediate checkpoints are used to evaluate the safety of the generation trajectory. The discretized path is regarded as safe if:

$$\min_{1 \leq r \leq R} \Delta_c(\boldsymbol{z}_r) \geq m_c^p, \tag{19}$$

$$\min_{1 \leq r \leq R} p_c(\boldsymbol{z}_r) \geq \eta_p, \qquad \max_{1 \leq r \leq R} \mathcal{H}\left(\boldsymbol{p}(\boldsymbol{z}_r)\right) \leq \xi_p. \tag{20}$$

If a candidate fails the path-level screening, BARGO adaptively shortens the interpolation length according to $t \leftarrow \beta t$, where $0 < \beta < 1$, and re-evaluates the candidate and its intermediate checkpoints. The candidate is rejected if it remains unsafe after $A_{\max}$ attempts. This mechanism treats reliability as a property of the entire generation trajectory rather than only the final synthesized point.

3.6 GEOMETRY-AWARE TOPOLOGY COORDINATED MODELING

A synthetic representation cannot participate in graph message-passing without compatible topology. For an accepted node $\widetilde{v}$ interpolated from v_i and v_j, BARGO first constructs a candidate-neighbor set as:

$$\mathcal{Q}_{\widetilde{v}} = \overline{\mathcal{N}}(i) \cup \overline{\mathcal{N}}(j) \cup \mathrm{kNN}_{\mathcal{M}}(\widetilde{\boldsymbol{h}}), \tag{21}$$

where $\overline{\mathcal{N}}(i) = \mathcal{N}(i) \cup \{v_i\}$ and $\mathrm{kNN}_{\mathcal{M}}(\widetilde{\boldsymbol{h}})$ contains the nearest real nodes to $\widetilde{v}$ under the geodesic distance.

For each $v_k \in \mathcal{Q}_{\widetilde{v}}$, the edge generator predicts its connection probability as:

$$e_{\widetilde{v}k} = \text{sigmoid}\left(\text{MLP}\left(\left[d_{\mathcal{M}}(\widetilde{\boldsymbol{h}}, \boldsymbol{h}_k) \| \kappa_k\right]\right)\right), \tag{22}$$

where $\|$ denotes feature concatenation, κ_k is the local graph curvature of v_k.

The target degree of the synthetic node is calibrated according to its source nodes:

$$\widetilde{d} = \min\left\{|\mathcal{Q}_{\widetilde{v}}|, \text{clip}\left(\text{round}\left((1-t)d_i + td_j\right), d_{\min}, d_{\max}\right)\right\}. \tag{23}$$

The proposed BARGO samples the $\widetilde{d}$ candidates with the largest edge probabilities and adds the corresponding edges to the augmented graph.

The edge generator is trained by re-constructing real graph connectivity:

$$\mathcal{L}_{\text{edge}} = -\frac{1}{|\boldsymbol{\mathcal{E}}^+|}\sum_{(u,v)\in\boldsymbol{\mathcal{E}}^+}\log e_{uv} - \lambda^- \frac{1}{|\boldsymbol{\mathcal{E}}^-|}\sum_{(u,v)\in\boldsymbol{\mathcal{E}}^-}\log(1-e_{uv}), \tag{24}$$

where $\mathcal{E}^+$ contains sampled real edges, $\mathcal{E}^-$ contains sampled non-edges, and λ^- balances the positive and negative re-construction terms.

Notably, the edge generator is optimized independently.

3.7 Optimization

The optimization objective on labeled samples is defined as:

$$\mathcal{L}_{\text{real}} = -\frac{1}{|\boldsymbol{\mathcal{V}}^L|}\sum_{v_i\in\boldsymbol{\mathcal{V}}^L}\log p_{i,y_i}. \tag{25}$$

The optimization objective on synthetic nodes is defined as:

$$\mathcal{L}_{\text{syn}} = -\frac{1}{|\boldsymbol{\mathcal{V}}^S|}\sum_{v_i\in\boldsymbol{\mathcal{V}}^S}\log p_{i,y_i}, \tag{26}$$

where $\boldsymbol{\mathcal{V}}^S$ denotes the set of accepted synthetic nodes.

To explicitly maintain boundary safety during joint optimization, we employ

$$\mathcal{L}_{\text{safe}} = \frac{1}{|\boldsymbol{\mathcal{V}}^S|}\sum_{v_i\in\boldsymbol{\mathcal{V}}^S}\left\{\left[m_{y_i} - \Delta_{y_i}(\boldsymbol{h}_i)\right]_+ + \frac{1}{\mathcal{R}}\sum_{r=1}^{\mathcal{R}}\left[m_{y_i}^{\text{path}} - \Delta_{y_i}(\boldsymbol{z}_{i,r})\right]_+\right\}, \tag{27}$$

where $[x]_+ = \max(0, x)$.

The overall objective is:

$$\mathcal{L} = \mathcal{L}_{\text{real}} + \lambda_1\mathcal{L}_{\text{syn}} + \lambda_2\mathcal{L}_{\text{safe}}. \tag{28}$$

4 Experiments

4.1 Experimental Settings

Datasets. We evaluate the proposed BARGO on six datasets, including three citation datasets (Cora, Citeseer, and PubMed) (Yang et al., 2016), two co-purchase datasets (Photo and Computers) and one co-author dataset (CS) (Shchur et al., 2018). Detailed statistics and additional experiments on three naturally imbalanced datasets are provided in Appendix A.1 to further evaluate the generalization of the proposed method.

Baselines. We compare BARGO with twelve competitive methods for class-imbalance issues. They are classified into two categories, including loss modifying methods (Reweight (Japkowicz & Stephen, 2002), PC Softmax (Hong et al., 2021), CB Loss (Cui et al., 2019), Focal Loss (Lin et al.,

2017), ReNode (Chen et al., 2021), and TAM (Song et al., 2022)), generative methods (Upsample, GraphSMOTE (Zhao et al., 2021), GraphENS (Park et al., 2021), GraphSHA (Li et al., 2023), and CDCGAN (Liu et al., 2025a)), and BAT (Liu et al., 2024) that addresses the class-imbalanced problem by topological augmentation.

Imbalance Settings. we validate the effectiveness of our proposed BARGO in manually imbalanced scenarios. Specifically, we conduct experiments under long-tailed class-imbalance settings on the citation datasets (Cora-LT, Citeseer-LT, and Pubmed-LT) and step class-imbalance settings on the co-purchase datasets (Photo-ST, Computers-ST) and the co-author dataset (CS-ST), respectively. In the long-tailed class-imbalance settings, the labeled nodes in the training set are removed until the imbalance ratio $\rho = 100$. In the step class-imbalance settings, half of the classes are majority classes and share an identical number of training samples, while the other half are minority ones and also hold the same number of training samples. The imbalance ratio is set to 20 in the step class-imbalance settings.

4.2 Experimental Results

Table 1: The experimental results (±std) on citation datasets of long-tailed class-imbalance settings. The reported results are the mean values of 10 runs, where the best result is highlighted in **bold**, and the runner-up is highlighted in underline.

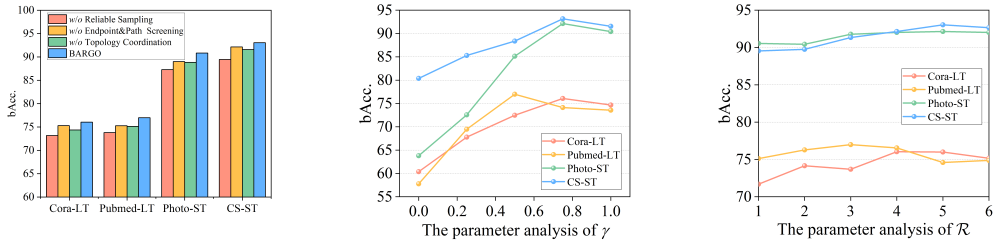
	Dataset	Cora-LT			Citeseer-LT			Pubmed-LT		
	ρ=100	Acc.	bAcc.	F1	Acc.	bAcc.	F1	Acc.	bAcc.	F1
GCN	Vanilla	72.02±0.50	59.42±0.74	59.23±1.02	51.40±0.44	44.64±0.42	37.82±0.67	51.58±0.60	42.11±0.48	34.73±0.71
	Re-weight	78.42±0.10	72.66±0.17	73.75±0.15	63.61±0.22	56.80±0.20	55.18±0.18	77.02±0.14	72.45±0.17	72.12±0.15
	PC Softmax	77.30±0.13	72.08±0.30	71.65±0.34	62.15±0.45	59.08±0.28	58.13±0.31	74.36±0.62	72.59±0.34	71.79±0.50
	CB Loss	77.97±0.19	72.70±0.28	73.17±0.22	61.47±0.51	55.18±0.52	53.47±0.65	76.57±0.19	72.16±0.18	72.84±0.19
	Focal Loss	78.43±0.19	73.17±0.23	73.76±0.20	59.66±0.38	53.39±0.33	51.80±0.39	75.67±0.20	71.34±0.24	72.03±0.21
	ReNode	78.93±0.13	73.13±0.17	74.46±0.16	62.39±0.31	55.62±0.27	54.05±0.24	76.00±0.16	70.68±0.15	71.41±0.15
	TAM	77.30±0.23	72.10±0.29	72.25±0.29	63.40±0.34	57.15±0.35	55.68±0.40	78.07±0.15	72.63±0.23	72.96±0.22
	Upsample	75.52±0.11	66.68±0.14	68.35±0.15	55.05±0.11	48.41±0.11	45.22±0.14	71.58±0.06	63.79±0.06	64.62±0.07
	GraphSMOTE	75.44±0.43	68.99±0.51	70.41±0.52	56.58±0.29	50.39±0.28	47.96±0.33	74.62±0.08	69.53±0.10	71.18±0.09
	GraphENS	76.15±0.24	71.16±0.40	70.85±0.49	63.14±0.35	56.92±0.37	55.54±0.41	77.11±0.11	71.89±0.15	72.71±0.14
	GraphSHA	79.90±0.29	74.62±0.35	75.74±0.32	64.50±0.41	59.04±0.34	59.16±0.21	79.20±0.13	74.46±0.17	75.24±0.27
	BAT	79.13±0.28	75.07±0.22	74.83±0.26	65.05±0.31	60.01±0.28	59.52±0.44	79.31±0.30	75.13±0.29	76.03±0.27
	CDCGAN	79.69±0.47	74.37±0.67	74.52±0.56	64.41±0.39	59.82±0.41	59.37±0.39	78.27±0.66	73.85±0.50	74.96±0.53
	BARGO	**80.77**±0.32	**76.03**±0.55	**76.37**±0.44	**65.59**±0.36	**60.20**±0.18	**59.61**±0.27	**80.88**±0.35	**76.98**±0.51	**77.23**±0.46

Table 2: The experimental results (±std) on co-purchase and co-author datasets in step class-imbalance settings.

	Dataset	Photot-ST			Computers-ST			CS-ST		
	ρ=20	Acc.	bAcc.	F1	Acc.	bAcc.	F1	Acc.	bAcc.	F1
GraphSAGE	Vanilla	59.19±2.32	59.98±1.82	47.11±2.95	63.88±0.05	46.96±0.04	30.08±0.11	74.72±0.33	79.62±0.28	65.11±0.51
	Re-weight	84.85±0.23	87.30±0.24	82.89±0.16	83.59±0.31	87.91±0.10	77.59±0.41	90.82±0.29	90.77±0.36	75.36±0.45
	PC Softmax	86.16±0.13	86.93±0.15	83.55±0.09	81.38±0.17	80.50±0.76	72.30±0.57	91.98±0.21	91.75±0.29	77.89±0.62
	CB Loss	83.02±0.29	85.79±0.21	80.48±0.31	83.75±0.21	87.38±0.10	77.08±0.20	90.52±0.55	90.57±0.19	79.66±0.88
	Focal Loss	82.58±0.39	85.42±0.32	79.28±0.35	82.56±0.22	87.38±0.08	76.53±0.15	90.02±0.18	90.00±0.23	79.53±0.15
	ReNode	84.83±0.15	86.43±0.20	81.85±0.22	81.29±0.34	87.33±0.17	76.60±0.28	91.07±0.36	91.15±0.44	80.29±0.51
	TAM	87.61±0.13	89.17±0.17	85.74±0.16	80.31±0.52	86.74±0.22	76.96±0.59	92.37±0.13	92.42±0.28	79.19±0.24
	Upsample	82.20±0.34	84.86±0.08	79.38±0.24	82.99±0.24	87.02±0.09	77.10±0.33	87.61±0.17	87.93±0.27	76.52±0.29
	GraphSMOTE	80.21±0.27	84.68±0.31	79.05±0.38	83.62±0.25	88.15±0.21	76.02±0.30	86.62±0.15	85.29±0.19	71.11±0.20
	GraphENS	88.02±0.09	90.55±0.11	86.70±0.10	83.28±0.38	88.54±0.10	76.77±0.52	92.13±0.16	92.59±0.24	78.27±0.28
	GraphSHA	89.14±0.22	90.60±0.10	87.25±0.18	84.21±0.50	89.49±0.10	77.93±0.91	92.65±0.08	92.71±0.12	79.55±0.16
	BAT	85.46±0.26	88.57±0.45	86.62±0.33	81.19±0.36	86.44±0.27	76.70±0.57	92.09±0.55	92.14±0.41	79.40±0.72
	CDCGAN	87.51±1.07	87.41±0.77	85.60±0.85	80.24±0.82	87.26±0.93	76.49±0.65	91.83±0.57	91.76±0.55	78.55±0.78
	BARGO	**90.83**±0.29	**92.14**±0.30	**89.42**±0.31	**86.65**±0.42	**89.75**±0.28	**78.69**±0.35	**93.33**±0.21	**93.04**±0.28	**80.55**±0.17

Experimental Results on Long-Tailed and Step Imbalance Issues. In this section, we conduct experiments under long-tailed class-imbalance settings on the citation datasets (Cora-LT, Citeseer-LT, Pubmed-LT) and step class-imbalance settings on the co-purchase and co-authors datasets (Photo-ST, Computers-ST, CS-ST). Specifically, we employ GCN as the backbone under the long-tailed class-imbalance settings, as well as leverage GraphSAGE as the backbone under the step class-imbalance settings. For evaluation, we use Accuracy (Acc.), balanced Accuracy (bAcc.), and macro F1 score (F1) as the metrics. The experimental results are shown in Table 1 and Table 2. We have the

following key observations. 1) The proposed method achieves superior performance compared to the competitive methods (i.e., loss modifying methods and oversampling methods), demonstrating the superiority of boundary-aware geodesic oversampling. 2) The proposed BARGO performs well on both the long-tailed and step class-imbalance settings, exhibiting the outstanding versatility of our proposed BARGO. 3) The proposed method yields a consistent improvement of performance across various datasets, proving the generalizability of our proposed BARGO.



(a) Ablation study. (b) Parameter analysis of γ. (c) Parameter analysis of $\mathcal{R}$.

Figure 3: The experimental results of ablation study and parameter analysis.

Ablation Study. In this section, we conduct ablation studies to verify the contribution of each component in the proposed method. Specifically, we remove reliable sampling, endpoint&path screening, and topology coordination separately. We conduct experiments on Cora-LT, Pubmed-LT, Photo-ST and CS-ST datasets. The experimental results are shown in Figure 3a. It can be observed that the removal of reliable sampling, endpoint&path screening, or topology coordination consistently reduces performance compared to the full model. The noticeable decline occurs when reliable sampling is removed, indicating that selecting reliable minority seeds and partners is important to prevent unreliable samples from being amplified during oversampling. The decline caused by the removal of endpoint&path screening supports the demand to check both the synthesized node and its synthesized trajectory for class consistency. Furthermore, topology coordination also contributes across all four datasets, demonstrating the importance of geometry-topology consistency for synthesized samples.

Parameter Analysis. In this section, we investigate the performance of the proposed method under key hyper-parameter settings. Specifically, γ controls the number of synthesized samples and $\mathcal{R}$ controls the number of geodesic checkpoints. We conduct experiments on Cora-LT, Pubmed-LT, Photo-ST and CS-ST datasets. The experimental results are shown in Figure 3b and Figure 3c. It can be observed that as γ increases from 0, demonstrating the benefit of adding minority samples. The performance of model peaks around $\gamma = 0.75$ on Cora-LT, Photo-ST, and CS-ST, while Pubmed-LT peaks around $\gamma = 0.5$. A further increase of γ to 1 does not improve performance and slightly reduces it. This suggests that moderate oversampling provides useful minority supervision, whereas unrestrained inclusion introduces redundant or unreliable synthesized nodes. For the parameter $\mathcal{R}$, increasing it from 1 generally improves the performance, as additional checkpoints provide a more thorough assessment of whether a geodesic path crosses an ambiguous region. The curves then fluctuate within a narrow range with their best results occurring. A further increase of $\mathcal{R}$ does not offer a consistent gain. Thus, a small number of checkpoints appears sufficient to capture most of the benefit of path-level screening without unnecessarily increasing its computational cost.

5 Conclusion

In this paper, we proposed BARGO, a boundary-aware Riemannian geodesic oversampling method for class-imbalanced graph learning. It samples reliable minority nodes by geometric margins and neighborhood consistency, then synthesizes samples along manifold geodesics. Furthermore, it introduces endpoint and path-level screening to reduce the risk of crossing category boundaries, while topology modeling and degree calibration provide compatible edges for synthesized nodes. The experimental results on long-tailed and step-imbalanced settings demonstrate improvements in the task of node classification over competitive methods. This work reveals that effective oversampling on graphs requires attention to the geometry of synthesized representations, the safety of their generation paths, and the consistency of their topology.

AI USE STATEMENT

During the preparation of this manuscript, the authors used a large language model for language polishing and grammatical refinement of author-drafted text. The AI tool was not used to generate core research ideas, model designs, experimental schemes or conclusions. All AI-suggested texts were fully reviewed, revised and verified manually by the authors. The authors take full responsibility for all content in this paper.

REFERENCES

Gregor Bachmann, Gary Bécigneul, and Octavian Ganea. Constant Curvature Graph Convolutional Networks. In *International Conference on Machine Learning*, pp. 486–496, 2020.

Ines Chami, Zhitao Ying, Christopher Ré, and Jure Leskovec. Hyperbolic Graph Convolutional Neural Networks. *Advances in Neural Information Processing Systems*, 32:1–12, 2019.

Deli Chen, Yankai Lin, Guangxiang Zhao, Xuancheng Ren, Peng Li, Jie Zhou, and Xu Sun. Topology-Imbalance Learning for Semi-Supervised Node Classification. *Advances in Neural Information Processing Systems*, 34:29885–29897, 2021.

Dongyue Chen, Zongxia Xie, Ruonan Liu, Wenlong Yu, Qinghua Hu, Xianling Li, and Steven X Ding. Bayesian Hierarchical Graph Neural Networks with Uncertainty Feedback for Trustworthy Fault Diagnosis of Industrial Processes. *IEEE Transactions on Neural Networks and Learning Systems*, pp. 18635–18648, 2023a.

Xiaohua Chen, Yucan Zhou, Dayan Wu, Chule Yang, Bo Li, Qinghua Hu, and Weiping Wang. AREA: Adaptive Reweighting via Effective Area for Long-Tailed Classification. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 19277–19287, 2023b.

Yin Cui, Menglin Jia, Tsung-Yi Lin, Yang Song, and Serge Belongie. Class-Balanced Loss Based on Effective Number of Samples. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9260–9269, 2019.

Yanhong Fei, Yingjie Liu, Chentao Jia, Zhengyu Li, Xian Wei, and Mingsong Chen. A survey of geometric optimization for deep learning: From euclidean space to riemannian manifold. *ACM Computing Surveys*, 57(5):1–37, 2025.

Youngkyu Hong, Seungju Han, Kwanghee Choi, Seokjun Seo, Beomsu Kim, and Buru Chang. Disentangling Label Distribution for Long-Tailed Visual Recognition. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 6626–6636, 2021.

Nathalie Japkowicz and Shaju Stephen. The Class Imbalance Problem: A Systematic Study. *Intelligent Data Analysis*, 6(5):429–449, 2002.

Wei Ju, Siyu Yi, Yifan Wang, Zhiping Xiao, Zhengyang Mao, Hourun Li, Yiyang Gu, Yifang Qin, Nan Yin, Senzhang Wang, et al. A Survey of Graph Neural Networks in Real World: Imbalance, Noise, Privacy and OOD Challenges. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 3036–3055, 2025.

Ranran Li, Zhaowei Liu, Yuanqing Ma, Dong Yang, and Shuaijie Sun. Internet Financial Fraud Detection Based on Graph Learning. *IEEE Transactions on Computational Social Systems*, 10 (3):1394–1401, 2022.

Wenzhi Li, Changdong Wang, Hui Xiong, and Jianhuang Lai. GraphSHA: Synthesizing Harder Samples for Class-Imbalanced Node Classification. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 1328–1340, 2023.

Zhixun Li, Yushun Dong, Qiang Liu, and Jeffrey Xu Yu. Rethinking Fair Graph Neural Networks from Re-Balancing. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 1736–1745, 2024.

Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. Focal Loss for Dense Object Detection. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 2980–2988, 2017.

Bojia Liu, Conghui Zheng, Fuhui Sun, Xiaoyan Wang, and Li Pan. CDCGAN: Class Distribution-Aware Conditional GAN-Based Minority Augmentation for Imbalanced Node Classification. *Neural Networks*, 183:106933, 2025a.

Qi Liu, Maximilian Nickel, and Douwe Kiela. Hyperbolic Graph Neural Networks. *Advances in neural information processing systems*, 32:1–12, 2019.

Zemin Liu, Yuan Li, Nan Chen, Qian Wang, Bryan Hooi, and Bingsheng He. A Survey of Imbalanced Learning on Graphs: Problems, Techniques, and Future Directions. *IEEE Transactions on Knowledge and Data Engineering*, 37(6):3132–3152, 2025b.

Zhining Liu, Ruizhong Qiu, Zhichen Zeng, Hyunsik Yoo, David Zhou, Zhe Xu, Yada Zhu, Kommy Weldemariam, Jingrui He, and Hanghang Tong. Class-Imbalanced Graph Learning without Class Rebalancing. In *International Conference on Machine Learning*, pp. 31747–31772, 2024.

Jiaan Luo, Feng Hong, Jiangchao Yao, Bo Han, Ya Zhang, and Yanfeng Wang. Revive Re-Weighting in Imbalanced Learning by Density Ratio Estimation. *Advances in Neural Information Processing Systems*, 37:79909–79934, 2024.

Joonhyung Park, Jaeyun Song, and Eunho Yang. GraphENS: Neighbor-Aware Ego Network Synthesis for Class-Imbalanced Node Classification. In *International Conference on Learning Representations*, pp. 1–20, 2021.

Hanyu Peng, Weiguo Pian, Mingming Sun, and Ping Li. Dynamic Re-Weighting for Long-Tailed Semi-Supervised Learning. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 6464–6474, 2023.

Wei Peng, Tuomas Varanka, Abdelrahman Mostafa, Henglin Shi, and Guoying Zhao. Hyperbolic Deep Neural Networks: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(12):10023–10044, 2021.

Saeed Rahmani, Asiye Baghbani, Nizar Bouguila, and Zachary Patterson. Graph Neural Networks for Intelligent Transportation Systems: A Survey. *IEEE Transactions on Information Forensics and Security*, 24(8):8846–8885, 2023.

Oleksandr Shchur, Maximilian Mumme, Aleksandar Bojchevski, and Stephan Günnemann. Pitfalls of Graph Neural Network Evaluation. *CoRR*, pp. 1–11, 2018.

Jaeyun Song, Joonhyung Park, and Eunho Yang. TAM: Topology-Aware Margin Loss for Class-Imbalanced Node Classification. In *International Conference on Machine Learning*, pp. 20369–20383, 2022.

Li Sun, Zhongbao Zhang, Junda Ye, Hao Peng, Jiawei Zhang, Sen Su, and Philip S Yu. A Self-Supervised Mixed-Curvature Graph Neural Network. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pp. 4146–4155, 2022.

Xinyue Wang, Yilin Lyu, and Liping Jing. Deep Generative Model for Robust Imbalance Classification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 14124–14133, 2020.

Xinyue Wang, Liping Jing, Yilin Lyu, Mingzhe Guo, Jiaqi Wang, Huafeng Liu, Jian Yu, and Tieyong Zeng. Deep Generative Mixture Model for Robust Imbalance Classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(3):2897–2912, 2022.

Zitai Wang, Qianqian Xu, Zhiyong Yang, Yuan He, Xiaochun Cao, and Qingming Huang. A Unified Generalization Analysis of Re-Weighting and Logit-Adjustment for Imbalanced Learning. *Advances in Neural Information Processing Systems*, 36:1–14, 2024.

Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, and S Yu Philip. A Comprehensive Survey on Graph Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems*, 32(1):4–24, 2020.

Guo-Sen Xie, Zheng Zhang, Guoshuai Liu, Fan Zhu, Li Liu, Ling Shao, and Xuelong Li. Generalized Zero-Shot Learning with Multiple Graph Adaptive Generative Networks. *IEEE Transactions on Neural Networks and Learning Systems*, 33(7):2903–2915, 2021.

Yu Xie, Yanfeng Liang, Maoguo Gong, A Kai Qin, Yew-Soon Ong, and Tiantian He. Semi-Supervised Graph Neural Networks for Graph Classification. *IEEE Transactions on Cybernetics*, 53(10):6222–6235, 2022.

Xihong Yang, Yiqi Wang, Yue Liu, Yi Wen, Lingyuan Meng, Sihang Zhou, Xinwang Liu, and En Zhu. Mixed Graph Contrastive Network for Semi-Supervised Node Classification. *ACM Transactions on Knowledge Discovery from Data*, pp. 1–9, 2024.

Zhilin Yang, William Cohen, and Ruslan Salakhudinov. Revisiting Semi-Supervised Learning with Graph Embeddings. In *International Conference on Machine Learning*, pp. 40–48, 2016.

Liang Zeng, Lanqing Li, Ziqi Gao, Peilin Zhao, and Jian Li. IMGCL: Revisiting Graph Contrastive Learning on Imbalanced Node Classification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pp. 11138–11146, 2023.

Muhan Zhang and Yixin Chen. Link Prediction Based on Graph Neural Networks. *Advances in Neural Information Processing Systems*, 31:1–11, 2018.

Yiding Zhang, Xiao Wang, Chuan Shi, Nian Liu, and Guojie Song. Lorentzian Graph Convolutional Networks. In *Proceedings of the web conference 2021*, pp. 1249–1261, 2021.

Tianxiang Zhao, Xiang Zhang, and Suhang Wang. GraphSMOTE: Imbalanced Node Classification on Graphs with Graph Neural Networks. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*, pp. 833–841, 2021.

A APPENDIX

A.1 DATASETS AND BASELINES

We use six widely-used datasets to evaluate the effectiveness of the proposed BARGO. These datasets contain three citation datasets (Cora, Citeseer, and Pubmed), two co-purchase datasets (Amazon-Photo and Amazon-Computers), and one co-author datasets. They are briefly described as follows and summarized in Table 3.

Table 3: The brief description of datasets.

Datasets	Nodes	Features	Edges	Classes	Imbalance Ratio
Cora	2,708	1,433	10,566	7	100
Citeseer	3,327	3,703	9,104	6	100
Pubmed	19,717	500	88,648	3	100
Photo	7,650	745	119,081	8	20
Computers	13,752	767	245,861	10	20
CS	18,333	6,805	81,894	15	20

Citation Dataset. The Cora, Citeseer, and Pubmed are three citation networks, where nodes represent papers and edges denote citation relations. In Cora and Citeseer datasets, each node is delineated by a 0/1-valued word vector that indicates the presence or absence of specific words from the dictionary. In Pubmed dataset, each node is described by a TF/IDF weighted word vector from the dictionary. The nodes are categorized by their related research area for the three datasets. Analogously, the ogbn-arXiv is a large-scale citation network induced from arXiv.

Co-purchase Dataset. The Amazon-Photo (Photo) and Amazon-Computers (Computers) are segments of the Amazon co-purchase graph, where nodes indicate goods and edges represent that two goods are frequently bought together. The features of a node are bag-of-words encoded product views. The class labels are given by the product category.

Co-author Dataset. The CS dataset represents an academic collaboration network in the field of computer science, where nodes correspond to authors and edges denote co-author relations. An edge is established between two authors if they have co-authored at least one publication. The node features are constructed using a bag-of-words model derived from the titles and abstracts of each author's publication. The node label indicates the primary research subfield of the author.

We compare the proposed GraphCALH with twelve competitive methods for class-imbalance issues. They are classified into two categories, including loss modifying methods (Re-weight Japkowicz & Stephen (2002), PC Softmax Hong et al. (2021), CB Loss **?**, Focal Loss Lin et al. (2017), ReNode Chen et al. (2021), and TAM Song et al. (2022)), generative methods (Upsample, GraphSMOTE Zhao et al. (2021), GraphENS Park et al. (2021), GraphSHA Li et al. (2023), and CDCGAN Liu et al. (2025a)), and BAT Liu et al. (2024) that addresses the class-imbalanced problem by topological augmentation. The details of those methods are introduced below.

Re-weight (Japkowicz & Stephen, 2002). It distributes different weights to samples of various categories, fine-tuning the contribution of each category in loss optimization, thereby compelling the model to pay more attention to the minority samples during training.

PC Softmax (Hong et al., 2021). It introduces an additional constraint into the Softmax loss function to ensure that the model considers the quantity distribution of samples among various categories.

CB Loss (?). It automatically adjusts the loss weights for each category, ensuring the model focuses on the minority categories during training without overlooking the majority.

Focal Loss (Lin et al., 2017). It reduces the focus of the model on easily classified samples and concentrates more on difficult minority samples, thus improving the recognition ability of the model for minority classes.

ReNode (Chen et al., 2021). It designs a metric based on conflict detection to identify the topological position of nodes and provides a unified perspective to jointly analyze the quantity and topology imbalance by taking into account the influence of node shift.

TAM (Song et al., 2022). It reveals that compensating minor nodes, which deviate from class-wise connectivity patterns, are prone to induce false positives for majority. Thus, it adjusts node-wise margins via the extent of deviation from connectivity patterns.

Upsample. It replicates minority samples to increase their observations, thereby balancing the quantity distribution across various categories and enhancing the representation of the minority ones.

GraphSMOTE (Zhao et al., 2021). It extends previous oversampling methods for i.i.d. data to the graph setting by constructing an intermediate embedding space with a feature extractor and training an edge generator and a GNN-based node classifier simultaneously on top of that.

GraphENS (Park et al., 2021). It investigates that existing imbalance handling methods suffer from the problem of neighbor memorization. Thus, it synthesizes an ego network for minor classes with neighbor sampling and saliency-based node mixing.

GraphSHA (Li et al., 2023). It synthesizes harder minor samples to enlarge the decision boundaries of minor classes and avoid invading the subspaces of neighbor classes.

BAT (Liu et al., 2024). It addresses the issue of class-imbalanced graph learning from a topological perspective. Furthermore, it is a model-agnostic method that can complement other techniques.

CDCGAN (Liu et al., 2025a). It introduces a conditional generative adversarial network to synthesize diverse and distinguishable minority nodes based on the characteristics of class distribution.