MEASURE BEFORE YOU PROPAGATE: SOURCE-RISK-AWARE MESSAGE ALLOCATION FOR GRAPH AUTOENCODING

Anonymous authors
Paper under double-blind review

ABSTRACT

Graph Neural Networks learn how neighboring messages should be combined, yet the reliability of the message source is often left implicit in the resulting interaction weights. Consequently, a source may be structurally relevant to its receiver while still providing a weakly supported representation whose influence spreads through neighborhood aggregation. We propose **SEGMA**, a source-risk-aware framework that separates source-level risk from pairwise interaction strength through a measurement–control design. SEGMA first computes spectral entropy as a training-independent proxy for feature dispersion and identifies potentially high-risk sources using adaptive distributional statistics. It then assigns the estimated risk to source-directed messages and incorporates it as an additive energy penalty before neighborhood normalization, yielding a monotone reallocation with respect to each message's own source risk without removing nodes or modifying graph topology. We further provide a conditional one-step analysis showing that the aggregation mass assigned to high-risk sources appears explicitly in an upper bound on representation deviation under fixed normalized weights. Experiments on diverse real-world attributed graphs demonstrate strong link-prediction performance, while component ablations, controlled perturbations, risk-mass diagnostics, and source-selection analyses provide complementary evidence for the proposed measurement–control mechanism. These results establish source reliability as a complementary dimension for graph message passing.

1 INTRODUCTION

Graph Neural Networks (GNNs) have become a standard paradigm for learning from relational data and are widely used in social network analysis (Khemani et al., 2024; Sharma et al., 2024), knowledge graph reasoning (He et al., 2026; Wang et al., 2025b), and recommender systems (Jendal et al., 2025; Nguyen et al., 2025). Their effectiveness largely stems from *collective inference*: each node is represented not only by its own attributes, but also by information propagated from its relational context (Zhang et al., 2025a). This view, however, gives every node a dual role—it is simultaneously an object to be represented and a source that influences others. Modern message-passing operators can assign non-uniform weights according to topology, pairwise compatibility, or task relevance (Bo et al., 2021; Dai et al., 2022; Wang et al., 2025a; Zhou et al., 2025). Yet interaction compatibility and source reliability are not equivalent. A message can be structurally relevant to its receiver while still originating from a weakly supported, unstable, or locally inconsistent representation. The central question is therefore not only *which nodes should interact*, but also *how much influence a potentially high-risk source should be allowed to exert*.

This distinction matters because source-level uncertainty need not coincide with conventional graph corruption. In real-world graphs, nodes with sparse observations, rapidly changing behavior, or weakly grounded attributes may remain structurally normal and should not be removed (Jiang et al., 2025; Cui et al., 2025; Wei et al., 2025b; Zhang et al., 2024b). Nevertheless, once their messages receive substantial aggregation weight, their influence can extend beyond the originating node and accumulate through successive propagation steps (Chen et al., 2024; Wang et al., 2025b). Figure 1 illustrates this distinction. The top panel depicts reliability-agnostic propagation, in which a potentially high-risk source can retain substantial influence because its reliability is not modeled

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.

separately from interaction strength. The bottom panel shows the desired alternative: the graph structure is preserved, but the effective influence of the source is regulated before its message spreads through the neighborhood. This perspective reframes robustness as an *information-flow control problem*, rather than solely an input-denoising problem.

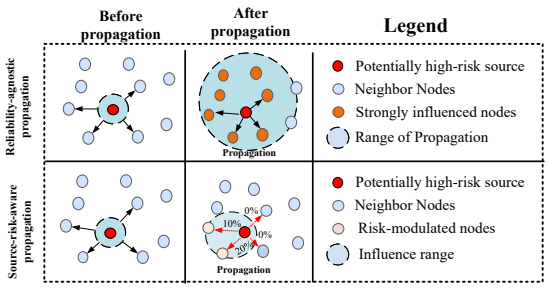


Figure 1: Reliability-agnostic vs. source-risk-aware propagation. **Top:** A potentially high-risk source may retain substantial influence without explicit reliability modeling. **Bottom:** Source-risk-aware propagation regulates its influence while preserving graph topology.

Existing studies address adjacent issues from different angles. Robust and topology-aware GNNs mainly target adversarial or noisy edges, structural perturbations, and heterophilous neighborhoods (Jo et al., 2023; Zhang et al., 2024a; 2023; Zügner & Günnemann, 2019). Attention-based and weighted propagation mechanisms adapt neighbor contributions according to learned compatibility or task-specific relevance (Bo et al., 2021; Wang et al., 2025a). Self-supervised graph autoencoders and masked learners improve representation quality through reconstruction, masking, or latent alignment (Zhao et al., 2025; Wei et al., 2023). These advances are effective, but they typically leave source reliability implicit in the learned interaction score. Consequently, they neither isolate a source-level risk signal from pairwise compatibility nor explicitly characterize the total normalized aggregation mass assigned to potentially high-risk sources. This missing separation can be consequential: when source risk is entangled with task relevance, a node may remain influential even if its representation is poorly aligned with the local evidence on which neighborhood aggregation relies. A principled solution should therefore connect two operations that are usually considered separately: *measuring* potentially risky sources and *controlling* their downstream influence.

Following the measurement–control perspective, we propose SEGMA, a source-reliability-guided message-allocation framework for self-supervised graph representation learning. SEGMA uses node-wise spectral entropy as a training-independent, scale-invariant proxy for feature dispersion, without assigning semantic meaning to individual frequencies or treating it as ground-truth unreliability. An adaptive IQR criterion identifies potentially high-risk sources from the upper tail of the score distribution. Their risks are then assigned to source-directed messages as additive energy penalties before neighborhood normalization, yielding monotone risk-conditioned reallocation while preserving graph connectivity and retaining useful information from all neighbors rather than removing selected sources. SEGMA is instantiated in a graph autoencoder and optimized by self-supervised structure reconstruction. We further show that risk mass—the normalized aggregation mass assigned to high-risk sources—appears explicitly in a conditional upper bound on one-step perturbation propagation, thereby linking source-level risk estimation, message allocation, and downstream propagation sensitivity while characterizing local sensitivity rather than global robustness.

Our contributions are summarized as follows:

- We distinguish *source reliability* from pairwise compatibility in graph message passing and formulate high-risk-source propagation as an explicit information-allocation problem. This perspective identifies risk mass as a concrete quantity connecting source-level uncertainty to propagation sensitivity.
- We propose SEGMA, a measurement–control framework that combines a training-independent feature-dispersion proxy with risk-aware message allocation. SEGMA regulates the effective influence of potentially high-risk sources while preserving the original topology and requiring no external supervision.
- We provide a conditional one-step perturbation analysis and conduct extensive link-prediction experiments across diverse real-world graphs. The results demonstrate strong and consistent predictive performance, while component ablations, controlled perturbations, and risk-mass diagnostics provide complementary evidence for the proposed measurement–control mechanism.

2 Related Work

Robust and Reliability-Aware Message Passing. Robust graph learning addresses unreliable information arising from adversarial perturbations, noisy edges, heterophily, and imperfect node attributes (Dai et al., 2022; Jo et al., 2023; Zügner & Günnemann, 2019; Zhang et al., 2024a). Attention- and structure-aware GNNs further reweight propagated messages according to pairwise compatibility, structural relevance, task importance, or node-specific signals (Bo et al., 2021; Wang et al., 2025a; Wei et al., 2023; Zhou et al., 2025; Mu et al., 2025). These methods demonstrate the value of adaptive propagation, but source reliability is typically inferred jointly with topology or task-dependent representations and remains entangled with interaction scores. As a result, a source may retain substantial normalized influence even when its representation is weakly supported by local evidence. SEGMA instead separates an input-level source-risk signal from pairwise interaction strength and explicitly uses it to regulate the normalized influence assigned to potentially high-risk sources during neighborhood message aggregation.

Graph Autoencoding and Masked Representation Learning. Graph autoencoders learn representations by reconstructing graph structure, node attributes, or latent targets. GAE/VGAE reconstruct adjacency from learned node embeddings (Kipf & Welling, 2016), while GraphMAE, GraphMAE2, SeeGera, and subsequent methods improve masking strategies, reconstruction objectives, latent targets, or structure-aware representation learning (Hou et al., 2022; 2023; Li et al., 2023; Wei et al., 2025a; Liu et al., 2025; Wang et al., 2024; He et al., 2024; Lim et al., 2024). Their primary focus is what information should be reconstructed and how reconstruction should be optimized. In contrast, SEGMA intervenes directly in encoder information flow by estimating source risk before training and modulating source-directed message contributions during aggregation, while preserving graph topology and requiring no reliability annotations.

3 Preliminaries and Problem Formulation

Collective Inference with Source-Level Risk. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{X})$ be an attributed graph with node features $\mathbf{x}_i \in \mathbb{R}^d$. A generic message-passing layer updates v_i as

$$\mathbf{h}_i^{(l+1)} = \sigma\left(\sum_{j \in \mathcal{N}(i)} \bar{\alpha}_{ij}^{(l)} \mathbf{W}^{(l)} \mathbf{h}_j^{(l)}\right), \tag{1}$$

where $\bar{\alpha}_{ij}^{(l)} \geq 0$ and $\sum_{j \in \mathcal{N}(i)} \bar{\alpha}_{ij}^{(l)} = 1$. Such weights typically encode interaction relevance without explicitly separating source reliability. We therefore study how normalized influence should be allocated to potentially high-risk sources.

High-Risk Sources and Risk Mass. Let $u_i = u(\mathbf{x}_i)$ be an input-level source-risk score and τ a threshold from its empirical distribution. We define

$$\mathcal{V}_r = \{v_i \in \mathcal{V} \mid u_i > \tau\} \tag{2}$$

as the set of potentially high-risk sources. Membership indicates elevated estimated risk rather than corruption, anomaly, or semantic error. For node v_i at layer l, its *risk mass* is

$$M_i^{(l)} = \sum_{j \in \mathcal{N}(i) \cap \mathcal{V}_r} \bar{\alpha}_{ij}^{(l)}, \tag{3}$$

the fraction of incoming aggregation influence assigned to high-risk sources. This quantity characterizes source-dependent propagation sensitivity under normalized message allocation without assuming that every such message is harmful.

Problem Statement. Given an attributed graph without source-reliability annotations, we aim to learn representations for self-supervised link prediction while regulating potentially high-risk sources. Specifically, we estimate source risk from input features and incorporate it into propagation without removing nodes or altering graph connectivity. The central question is: *how can source-level risk be estimated before training and translated into propagation-level control?*

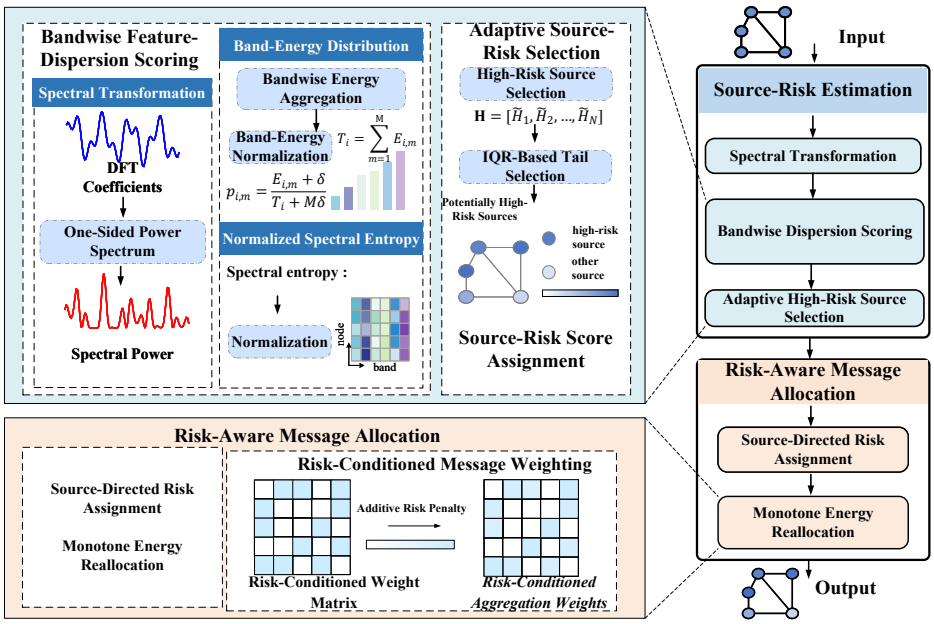


Figure 2: Overview of SEGMA under the measurement–control formulation. **Source-Risk Estimation** uses bandwise spectral entropy and an adaptive IQR criterion to identify potentially high-risk sources. **Risk-Aware Message Allocation** assigns source risk to directed messages and penalizes their energies before normalization, regulating aggregation without modifying graph topology. Representations are learned through self-supervised graph reconstruction.

4 METHODOLOGY

We present **SEGMA**, a source-reliability-guided message-allocation framework for self-supervised graph representation learning. SEGMA follows a measurement–control principle: before regulating information flow, it first constructs an operational estimate of which nodes may constitute high-risk message sources and then translates this estimate into source-directed control of normalized message contributions. As illustrated in Figure 2, the framework contains two coordinated components: *Source-Risk Estimation*, which computes a training-independent feature-dispersion score and selects potentially high-risk sources, and *Risk-Aware Message Allocation*, which conditions directed message weights on the estimated source risk without removing nodes or altering graph connectivity.

4.1 SOURCE-RISK ESTIMATION

Spectral Transformation. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{X})$ be an attributed graph with N nodes, where node v_i is associated with a feature vector $\mathbf{x}_i \in \mathbb{R}^d$. SEGMA uses spectral entropy as an operational proxy for feature dispersion under the observed feature coordinate system. It does not assign semantic meanings to individual frequency components or treat spectral entropy as a ground-truth annotation of source reliability. We first apply a discrete Fourier transform to each feature vector and retain its one-sided power spectrum:

$$\widehat{\mathbf{x}}_i = \mathcal{F}(\mathbf{x}_i), \qquad p_i(k) = |\widehat{x}_i(k)|^2, \quad k = 1, \ldots, K, \tag{4}$$

where $K = \lfloor d/2 \rfloor + 1$ for real-valued input features. The transformation is performed once before model training.

Bandwise Feature-Dispersion Scoring. To obtain a compact and controllable spectral description, we partition the one-sided spectrum into q non-overlapping bands $\{\mathcal{B}_m\}_{m=1}^{q}$. The energy in band m and its normalized distribution are defined as

$$E_{i,m} = \sum_{k \in \mathcal{B}_m} p_i(k), \qquad \pi_{i,m} = \frac{E_{i,m} + \epsilon}{\sum_{t=1}^{q} E_{i,t} + q\epsilon}, \tag{5}$$

where $\epsilon > 0$ prevents numerical instability. We then compute the normalized spectral-entropy score

$$H_i = -\sum_{m=1}^{q} \pi_{i,m} \log \pi_{i,m}, \qquad s_i = \frac{H_i}{\log q}. \tag{6}$$

The resulting score satisfies $s_i \in [0, 1]$. A larger value indicates that the feature energy is distributed more broadly across the predefined bands, whereas a smaller value indicates a more concentrated profile. Because the band-energy distribution is normalized, s_i is invariant to uniform scaling of $\mathbf{x}_i$ and remains fixed throughout optimization. However, the spectral proxy operates on the observed feature coordinate system, and arbitrary permutations of feature dimensions may alter the resulting scores. We therefore regard it as an operational estimator for attributed graphs with a fixed feature organization, rather than a permutation-invariant measure of intrinsic source reliability.

Adaptive High-Risk Source Selection. Source risk is interpreted relative to the empirical score distribution of each graph rather than through a universal threshold. Let Q_1 and Q_3 denote the first and third quartiles of $\mathcal{S} = \{s_i\}_{i=1}^{N}$, and let $\mathrm{IQR} = Q_3 - Q_1$. We define

$$\tau = Q_3 + \lambda \cdot \mathrm{IQR}, \qquad \mathcal{V}_r = \{v_i \in \mathcal{V} \mid s_i > \tau\}, \tag{7}$$

where λ controls the selectivity of upper-tail identification. Membership in $\mathcal{V}_r$ does not imply that a node is corrupted, anomalous, or semantically incorrect. It indicates only that the node receives an elevated score under the adopted feature-dispersion estimator and is consequently treated as a potentially high-risk message source.

4.2 Risk-Aware Message Allocation

Source-Directed Risk Assignment. Node-level risk affects collective inference only after it is associated with individual messages. Because the risk considered in this work concerns the information source, SEGMA assigns risk according to the direction of message transmission. For a message from source node v_j to receiver node v_i, we define

$$r_{j \to i} = \mathbb{I}(v_j \in \mathcal{V}_r)\, s_j. \tag{8}$$

For an undirected graph, an observed edge is represented by two directed messages, and each direction receives the score of its corresponding source. This formulation avoids symmetrically penalizing a low-risk message merely because its receiver has a high source-risk score.

Monotone Energy Reallocation. Let $\ell_{ij}^{(l)}$ denote the base interaction logit assigned to message $v_j \to v_i$ at layer l. Without source-risk conditioning, the positive unnormalized interaction mass is $\exp(\ell_{ij}^{(l)})$. SEGMA incorporates source risk through an additive energy penalty:

$$\ell_{ij}'^{(l)} = \ell_{ij}^{(l)} - \beta r_{j \to i}, \qquad \beta \geq 0, \tag{9}$$

and obtains the normalized aggregation weight

$$\overline{\alpha}_{ij}^{(l)} = \frac{\exp\left(\ell_{ij}^{(l)} - \beta r_{j \to i}\right)}{\sum_{k \in \mathcal{N}(i)} \exp\left(\ell_{ik}^{(l)} - \beta r_{k \to i}\right)}. \tag{10}$$

Equivalently, Equation equation 10 multiplies the positive unnormalized interaction mass by $\exp(-\beta r_{j \to i})$ before neighborhood normalization. This placement avoids the sign reversal that can arise when a gate smaller than one is directly multiplied by a negative logit. Conditioned on the remaining logits and risk scores, the weight of a message is monotone non-increasing with respect to its own source risk:

$$\frac{\partial \overline{\alpha}_{ij}^{(l)}}{\partial r_{j \to i}} = -\beta \overline{\alpha}_{ij}^{(l)} \left(1 - \overline{\alpha}_{ij}^{(l)}\right) \leq 0. \tag{11}$$

SEGMA therefore performs differentiable reallocation over the neighborhood simplex rather than deleting selected neighbors or assigning a fixed binary mask.

Risk-Mass-Aware Aggregation. After the risk-conditioned weights have been normalized, the risk mass received by node v_i is

$$M_i^{(l)} = \sum_{j \in \mathcal{N}(i) \cap \mathcal{V}_r} \overline{\alpha}_{ij}^{(l)}. \tag{12}$$

Risk mass measures the fraction of normalized incoming influence allocated to designated high-risk sources. It is an analytical and diagnostic quantity rather than an additional supervision signal. The representation update is

$$\mathbf{h}_i^{(l+1)} = \sigma\left(\sum_{j \in \mathcal{N}(i)} \overline{\alpha}_{ij}^{(l)} \mathbf{W}^{(l)} \mathbf{h}_j^{(l)}\right). \tag{13}$$

SEGMA does not assume that every message originating from $\mathcal{V}_r$ is harmful. Instead, it introduces an explicit mechanism through which a learned interaction score can be conditioned on a stable source-level risk signal.

4.3 Self-Supervised Graph Reconstruction

SEGMA is instantiated within a graph autoencoder so that source-risk-aware propagation can be optimized without reliability annotations. The encoder applies Equation equation 13 to obtain node embeddings $\mathbf{z}_i$. Using an inner-product link decoder, the probability of an edge between v_i and v_j is

$$\widehat{A}_{ij} = \text{sigmoid}\left(\mathbf{z}_i^\top \mathbf{z}_j\right). \tag{14}$$

Let $\Omega = \mathcal{E}^+ \cup \mathcal{E}^-$ contain observed training edges and sampled non-edges, with $A_{ij} \in \{0, 1\}$ denoting the corresponding target. The reconstruction objective is

$$\mathcal{L}_{\text{rec}} = -\sum_{(i,j) \in \Omega} \left[A_{ij} \log \widehat{A}_{ij} + (1 - A_{ij}) \log(1 - \widehat{A}_{ij})\right]. \tag{15}$$

The source-risk scores and selected source set remain fixed during training, whereas the encoder parameters and base interaction logits are optimized end-to-end through graph reconstruction.

4.4 Complexity Analysis

Source-risk estimation is performed once before training. The spectral transformation requires $\mathcal{O}(Nd \log d)$ time, while band aggregation and entropy computation require $\mathcal{O}(Nd)$. At each message-passing layer, source-risk assignment, energy penalization, and neighborhood normalization introduce $\mathcal{O}(E)$ operations, while feature transformation and sparse aggregation require $\mathcal{O}(Em)$ time for hidden dimension m. The overall time complexity of an L-layer encoder is therefore

$$\mathcal{O}(Nd \log d + LEm), \tag{16}$$

with an additional linear edge-level overhead that does not change the asymptotic order of standard sparse message passing. SEGMA preserves the original adjacency structure and introduces neither edge pruning nor auxiliary graph construction.

5 Risk-Mass Characterization

We provide a local characterization of how the aggregation mass assigned to high-risk sources affects one-step propagation sensitivity. The analysis is conditioned on fixed normalized weights and is intended to clarify the role of risk mass, rather than establish global robustness for a complete nonlinear GNN. Consider the message-passing layer

$$\mathbf{h}_i^{(l+1)} = \sigma\left(\sum_{j \in \mathcal{N}(i)} \bar{\alpha}_{ij}^{(l)} \mathbf{W}^{(l)} \mathbf{h}_j^{(l)}\right), \tag{17}$$

where $\bar{\alpha}_{ij}^{(l)} \geq 0$, $\sum_{j \in \mathcal{N}(i)} \bar{\alpha}_{ij}^{(l)} = 1$, and $\sigma(\cdot)$ is L_σ-Lipschitz. Recall that the risk mass of node v_i is $M_i^{(l)} = \sum_{j \in \mathcal{N}(i) \cap \mathcal{V}_r} \bar{\alpha}_{ij}^{(l)}$.

Proposition 1 (One-step sensitivity under risk mass)**.** *Let $\Delta\mathbf{h}_j^{(l)}$ be an additive perturbation supported on the designated high-risk set $\mathcal{V}_r$, such that $\Delta\mathbf{h}_j^{(l)} = \mathbf{0}$ for $v_j \notin \mathcal{V}_r$. Define $\delta_r^{(l)} = \max_{v_j \in \mathcal{V}_r} \|\Delta\mathbf{h}_j^{(l)}\|$. Conditioned on fixed aggregation weights, the resulting representation deviation satisfies*

$$\left\|\Delta\mathbf{h}_i^{(l+1)}\right\| \leq L_\sigma \left\|\mathbf{W}^{(l)}\right\| M_i^{(l)} \delta_r^{(l)}. \tag{18}$$

The result follows directly from the Lipschitz continuity of $\sigma(\cdot)$, the triangle inequality, and the definition of $M_i^{(l)}$; the proof is provided in **Appendix** A. Proposition 1 identifies risk mass as an explicit factor in the conditional upper bound on one-step representation deviation. Thus, under the same perturbation budget and transformation norm, a smaller risk mass yields a tighter sensitivity bound. This result does not establish the optimality of the adopted source-risk estimator, guarantee that the learned controller always decreases $M_i^{(l)}$, or characterize multi-layer global stability. Instead, it motivates risk-aware message allocation and provides a theoretical lens for the representation-deviation diagnostics in Section Experiments.

6 Experiments

6.1 Experimental Setup

Datasets and Task. We evaluate SEGMA on nine real-world attributed graphs covering citation networks (**Cora**, **CiteSeer**, and **PubMed**) (Sen et al., 2008), co-purchase networks (**Photo** and **Computers**) (Shchur et al., 2018), co-authorship networks (**CS** and **Physics**) (Shchur et al., 2018), and social networks (**Actor** and **BlogCatalog**) (Pei et al., 2020; Tang & Liu, 2009). Dataset statistics are provided in **Appendix** B.

Baselines. We compare SEGMA with representative self-supervised graph learning methods from several paradigms: reconstruction-based **GAE** and **VGAE** (Kipf & Welling, 2016); contrastive methods **MVGRL** (Hassani & Khasahmadi, 2020) and **GCA** (Zhu et al., 2021); masked graph autoencoders **GraphMAE** (Hou et al., 2022) and **GraphMAE2** (Hou et al., 2023); variational masking with **SeeGera** (Li et al., 2023); latent-target reconstruction with **GiGaMAE** (Shi et al., 2024); and teacher-guided graph autoencoding with **TED-GAE** (Zhang et al., 2025b); and attention-based link-prediction model with **BS-SubGNN** (Fang et al., 2026). We use official implementations whenever available and apply the same data splits and evaluation pairs to all methods.

Parameter Settings Following established link-prediction protocols (Kipf & Welling, 2016; Shi et al., 2024), we evaluate all methods using AUC and AP, with higher values indicating better performance. Results are reported as mean $\pm$ standard deviation over five runs with different random seeds. Baselines are implemented using their official repositories whenever available and evaluated under the same data splits, negative samples, and evaluation protocol. Their hyperparameters are selected according to validation performance within the ranges recommended by the original papers or official implementations. For SEGMA, source-risk scores are computed once from the input features using FFT-based spectral entropy and remain fixed during training. The backbone consists of a two-layer risk-aware GCN encoder followed by an inner-product link decoder.

6.2 Overall Link-Prediction Performance

Table 1 compares SEGMA with representative graph autoencoding, contrastive, and masked-learning baselines. SEGMA ranks first on both metrics for every dataset. The improvements are observed across datasets with markedly different sizes and feature spaces, suggesting that source-risk-aware message allocation is not tied to a single graph category. In particular, SEGMA substantially improves over the shared GAE backbone on Actor, BlogCatalog, CiteSeer, Photo, and Computers, indicating that its gains cannot be attributed solely to the reconstruction objective. The comparison with recent masked autoencoders further shows that modifying the information-flow process provides a complementary benefit to improving masking or reconstruction targets. We avoid attributing individual gains to unmeasured dataset properties; the following analyses instead examine whether the observed improvements are consistent with the proposed measurement–control mechanism.

Table 1: Link-prediction performance on nine attributed graphs. Results are reported as mean ± standard deviation over five runs. The best result in each column is highlighted in bold.

Metric	Method	Actor	BlogCatalog	CiteSeer	Computers	Cora	CS	Photo	Physics	PubMed
AUC	GAE (2016)	70.19±0.91	75.68±0.88	89.53±0.29	81.29±0.79	91.00±0.16	95.47±0.35	81.46±0.26	96.08±0.36	96.23±0.07
	VGAE (2016)	70.90±0.28	81.92±0.92	89.66±0.89	49.79±0.87	88.50±0.85	95.82±0.83	87.19±0.81	96.10±0.78	94.98±0.75
	MVGRL (ICML'20)	72.34±0.34	84.14±0.70	67.23±0.08	82.16±0.36	75.53±0.24	76.10±0.12	84.77±0.60	77.13±0.58	91.11±0.28
	GCA (NeurIPS'20)	77.48±0.84	86.20±0.52	87.01±0.10	90.79±0.40	89.96±0.46	87.93±0.24	91.42±0.12	90.63±0.90	90.21±0.38
	GraphMAE (KDD'22)	76.54±0.57	85.22±0.75	95.23±0.55	71.23±0.71	93.21±0.87	92.94±0.27	75.56±0.45	85.78±0.03	87.29±0.40
	GraphMAE2 (WWW'23)	76.58±0.69	85.52±0.08	95.26±0.26	72.02±0.51	93.34±0.38	92.90±0.82	73.06±0.50	89.51±0.46	90.25±0.40
	SeeGera (WWW'23)	74.30±0.47	86.18±0.60	88.29±0.08	96.49±0.36	93.91±0.57	95.46±0.50	97.28±0.65	OOM	87.50±0.54
	GiGaMAE (CIKM'24)	69.96±0.53	84.30±0.32	96.47±0.30	93.53±0.61	94.98±0.32	96.92±0.58	94.20±0.60	96.20±0.30	95.17±0.59
	TED-GAE (AAAI'25)	69.10±0.83	OOM	95.13±0.64	90.81±0.37	94.35±0.05	96.32±0.29	93.37±0.18	OOM	92.77±0.60
	BS-SubGNN (AAAI'26)	68.23±0.25	83.45±0.09	94.73±0.31	73.12±0.33	90.24±0.28	91.88±0.50	72.28±0.30	88.59±0.22	90.11±0.30
	SEGMA	**80.08±1.06**	**87.30±0.27**	**96.66±0.34**	**97.21±0.25**	**95.85±0.11**	**97.13±0.32**	**97.48±0.09**	**97.11±0.11**	**97.48±0.11**
AP	GAE (2016)	74.72±0.94	75.82±0.27	90.27±0.55	82.13±0.90	91.75±0.83	95.13±0.88	80.41±0.59	95.75±0.86	96.56±0.85
	VGAE (2016)	75.26±0.30	82.16±0.82	91.42±0.29	56.95±0.83	89.54±0.90	95.66±0.18	86.96±0.27	95.71±0.06	95.28±0.75
	MVGRL (ICML'20)	75.12±0.04	83.41±0.73	71.65±0.42	83.59±0.25	75.16±0.36	74.13±0.46	85.49±0.39	78.74±0.67	91.13±0.60
	GCA (NeurIPS'20)	76.28±0.65	86.03±0.36	88.51±0.45	90.34±0.84	88.72±0.08	86.89±0.60	92.52±0.68	91.23±0.50	92.35±0.57
	GraphMAE (KDD'22)	77.02±0.56	84.26±0.88	94.56±0.60	66.04±0.23	91.36±0.52	92.60±0.67	71.09±0.50	82.43±0.49	86.05±0.39
	GraphMAE2 (WWW'23)	77.16±0.47	84.31±0.66	94.54±0.16	67.10±0.48	91.48±0.37	92.64±0.42	70.09±0.38	85.43±0.40	90.45±0.30
	SeeGera (WWW'23)	77.31±0.79	85.97±0.32	91.11±0.36	96.35±0.19	94.25±0.45	92.79±0.34	96.72±0.18	OOM	87.07±0.33
	GiGaMAE (CIKM'24)	66.10±0.07	82.14±0.31	96.75±0.17	91.21±0.30	94.95±0.06	96.40±0.15	91.46±0.23	94.47±0.28	94.38±0.27
	TED-GAE (AAAI'25)	67.51±0.26	OOM	95.09±0.22	87.82±0.48	93.77±0.59	94.21±0.22	90.18±0.25	OOM	90.93±0.20
	BS-SubGNN (AAAI'26)	67.12±0.08	82.89±0.21	93.89±0.28	72.56±0.18	90.29±0.35	90.58±0.45	71.78±0.45	86.23±0.36	90.05±0.23
	SEGMA	**80.73±0.19**	**86.77±0.46**	**97.13±0.18**	**96.68±0.16**	**95.75±0.30**	**97.23±0.14**	**97.04±0.16**	**96.84±0.08**	**97.07±0.35**

6.3 Component Analysis

We examine whether source-risk estimation and propagation control provide complementary contributions. **w/o Risk Estimation** removes the spectral-entropy partition and applies no source-specific risk signal. **w/o Risk Control** retains the estimated source set but replaces risk-aware allocation with backbone propagation. **Hard Risk Control** assigns fixed attenuation to interactions from selected sources, while **Random Risk Assignment** selects the same number of sources uniformly at random.

Table 2: Ablation study in terms of AUC.

Variant	CiteSeer	CS	Physics
w/o Risk Estimation	95.89	95.96	96.04
w/o Risk Control	89.53	94.47	95.08
Hard Risk Control	94.98	95.05	95.97
Random Risk Assignment	95.13	96.03	95.11
SEGMA	**96.66**	**97.13**	**97.11**

As shown in Table 2, SEGMA performs best across all evaluated datasets. Removing risk-aware propagation causes the largest degradation, showing that source-risk estimation alone is insufficient without propagation-level control. Removing risk estimation also consistently reduces performance, indicating that the control module benefits from a data-dependent source signal rather than merely introducing another weighting layer. Neither fixed attenuation nor random source assignment matches the full model, supporting the complementarity of measurement and control without treating the selected nodes as ground-truth unreliable sources.

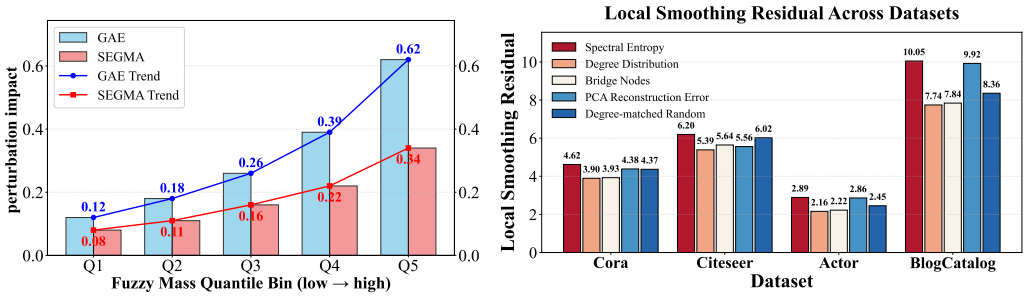
6.4 Does Risk Mass Characterize Propagation Sensitivity?

We examine whether risk mass characterizes one-step propagation sensitivity. After training, we freeze the encoder and inject feature perturbations only into selected high-risk sources. For each model, we compare clean and perturbed forward passes and measure the first-layer representation deviation of node v_i as

$$D_i = \left\| \mathbf{h}_i^{(1)}(\mathbf{X} + \Delta\mathbf{X}) - \mathbf{h}_i^{(1)}(\mathbf{X}) \right\|_2 . \tag{19}$$

For a matched comparison, we compute risk mass from baseline propagation weights, partition nodes into five exposure quantiles, and apply the same partition to GAE and SEGMA.

This experiment diagnoses conditional one-step sensitivity rather than global robustness. As shown in Figure 3 (a), GAE's representation deviation increases with risk mass, indicating stronger perturbation transmission under greater exposure to high-risk sources. SEGMA shows the same trend but consistently lower deviations, with larger differences in high-risk regions. This is consistent with Proposition 1 , which identifies risk mass as a factor in the conditional sensitivity bound. Thus,



(a) Risk-mass sensitivity (b) Source-risk selection

Figure 3: Mechanism diagnostics. (a) One-step representation deviation across risk-mass quantiles, where "Fuzzy mass" denotes $M_i^{(l)}$. (b) Local smoothness residual under different source-risk estimators with matched selection ratios.

risk mass serves as a propagation-level diagnostic without implying complete characterization or multi-layer stability.

6.5 What Does the Source-Risk Proxy Select?

Because ground-truth source-reliability annotations are unavailable, we evaluate spectral entropy through an aggregation-oriented diagnostic rather than as a measure of semantic correctness. We compare it with degree-based selection, betweenness centrality, PCA reconstruction error, and degree-matched random selection, with all strategies selecting the same number of nodes. For each selected node v_i, we compute the local smoothness residual

$$R_i = \left\| \mathbf{x}_i - \frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \mathbf{x}_j \right\|_2, \tag{20}$$

which measures feature disagreement with the immediate neighborhood. A larger residual indicates weaker alignment with the local smoothness assumption underlying neighborhood aggregation, rather than semantic incorrectness or corruption.

Figure 3(b) shows that spectral-entropy-based selection yields the largest local smoothness residual across datasets. Degree- and betweenness-based criteria capture structural prominence, while degree-matched random selection suggests that the difference is not explained by node degree. PCA reconstruction error provides a stronger feature-based alternative but remains below spectral entropy. These results do not establish that spectral entropy recovers an unobserved reliability ground truth; rather, they show that it preferentially selects nodes whose features are less aligned with neighborhood evidence, supporting its use as an operational source-risk proxy.

6.6 Additional Experiments

Appendix C provides additional analyses of **feature perturbations**, **structural corruption**, **hyperparameter sensitivity**, and **runtime**. Perturbations serve as diagnostic stress tests rather than simulations of semantic unreliability, while runtime results reflect empirical costs without implying edge pruning or reduced message-passing complexity.

7 Conclusion

We presented **SEGMA**, a source-risk-aware framework that separates source-level risk from pairwise interaction strength. SEGMA combines a training-independent feature-dispersion proxy with risk-aware message allocation without modifying graph topology. A conditional analysis links risk mass to one-step propagation sensitivity, while experiments demonstrate strong link-prediction performance and consistent mechanism-level evidence. These results highlight source reliability as a complementary dimension of message passing.

Reproducibility Statement

Implementation details and experimental protocols are provided in Section 6 and Appendix B–C. The proof of Proposition 1 is included in Appendix A, while dataset statistics, perturbation settings, hyperparameter analyses, and runtime evaluations are provided in the appendix.

AI Use Statement

Generative AI tools were used solely for language polishing and improving the clarity and readability of the manuscript. All methodological decisions, theoretical claims, experiments, and reported results were developed, conducted, and verified by the authors.

References

Deyu Bo, Xiao Wang, Chuan Shi, and Huawei Shen. Beyond low-frequency information in graph convolutional networks. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, pp. 3950–3957, 2021.

Hao Chen, Yuanchen Bei, Qijie Shen, Yue Xu, Sheng Zhou, Wenbing Huang, Feiran Huang, Senzhang Wang, and Xiao Huang. Macro graph neural networks for online billion-scale recommender systems. In Tat-Seng Chua, Chong-Wah Ngo, Ravi Kumar, Hady W. Lauw, and Roy Ka-Wei Lee (eds.), *Proceedings of the ACM on Web Conference 2024, WWW 2024, Singapore, May 13-17, 2024*, pp. 3598–3608. ACM, 2024. doi: 10.1145/3589334.3645517. URL https://doi.org/10.1145/3589334.3645517.

Ziqiang Cui, Yunpeng Weng, Xing Tang, Fuyuan Lyu, Dugang Liu, Xiuqiang He, and Chen Ma. Comprehending knowledge graphs with large language models for recommender systems. In Nicola Ferro, Maria Maistro, Gabriella Pasi, Omar Alonso, Andrew Trotman, and Suzan Verberne (eds.), *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2025, Padua, Italy, July 13-18, 2025*, pp. 1229–1239. ACM, 2025. doi: 10.1145/3726302.3729932. URL https://doi.org/10.1145/3726302.3729932.

Enyan Dai, Wei Jin, Hui Liu, and Suhang Wang. Towards robust graph neural networks for noisy graphs with sparse labels. In *Proceedings of the fifteenth ACM international conference on web search and data mining*, pp. 181–191, 2022.

Zhihong Fang, Shaolin Tan, Qiu Fang, Zhe Li, and Qing Gao. Subgraph encoding with bicentric sphere node labeling and pooling for link prediction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 14711–14719, 2026.

Kaveh Hassani and Amir Hosein Khasahmadi. Contrastive multi-view representation learning on graphs. In *International conference on machine learning*, pp. 4116–4126. PMLR, 2020.

Junwei He, Qianqian Xu, Yangbangyan Jiang, Zitai Wang, and Qingming Huang. ADA-GAD: anomaly-denoised autoencoders for graph anomaly detection. In Michael J. Wooldridge, Jennifer G. Dy, and Sriraam Natarajan (eds.), *Thirty-Eighth AAAI Conference on Artificial Intelligence, AAAI 2024, Thirty-Sixth Conference on Innovative Applications of Artificial Intelligence, IAAI 2024, Fourteenth Symposium on Educational Advances in Artificial Intelligence, EAAI 2014, February 20-27, 2024, Vancouver, Canada*, pp. 8481–8489. AAAI Press, 2024. doi: 10.1609/AAAI.V38I8.28691. URL https://doi.org/10.1609/aaai.v38i8.28691.

Tao He, Zerui Chen, Lizi Liao, Yixin Cao, Yuanxing Liu, Wei Tang, Xun Mao, Kai Lv, Ming Liu, and Bing Qin. Subgraph-centric multi-agent reinforcement learning for multi-hop knowledge graph reasoning. *IEEE Trans. Knowl. Data Eng.*, 38(2):1319–1333, 2026. doi: 10.1109/TKDE.2025.3631495. URL https://doi.org/10.1109/TKDE.2025.3631495.

Zhenyu Hou, Xiao Liu, Yukuo Cen, Yuxiao Dong, Hongxia Yang, Chunjie Wang, and Jie Tang. Graphmae: Self-supervised masked graph autoencoders. In *Proceedings of the 28th ACM SIGKDD conference on knowledge discovery and data mining*, pp. 594–604, 2022.

Zhenyu Hou, Yufei He, Yukuo Cen, Xiao Liu, Yuxiao Dong, Evgeny Kharlamov, and Jie Tang. Graphmae2: A decoding-enhanced masked self-supervised graph learner. In *Proceedings of the ACM web conference 2023*, pp. 737–746, 2023.

Theis E. Jendal, Matteo Lissandrini, Peter Dolog, and Katja Hose. The limits of graph samplers for training inductive recommender systems. *Proc. VLDB Endow.*, 18(8):2496–2504, 2025. doi: 10.14778/3742728.3742743. URL https://www.vldb.org/pvldb/vol18/p2496-jendal.pdf.

Linling Jiang, Wenchang Zhang, Hua Wang, and Fan Zhang. Robust memory-based graph neural networks for noisy and sparse graphs. *Engineering Applications of Artificial Intelligence*, 160: 111874, 2025.

Hyeonsoo Jo, Fanchen Bu, and Kijung Shin. Robust graph clustering via meta weighting for noisy graphs. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, pp. 1035–1044, 2023.

Bharti Khemani, Shruti Patil, Ketan Kotecha, and Sudeep Tanwar. A review of graph neural networks: concepts, architectures, techniques, challenges, datasets, applications, and future directions. *Journal of Big Data*, 11(1):18, 2024.

Thomas N Kipf and Max Welling. Variational graph auto-encoders. *arXiv preprint arXiv:1611.07308*, 2016.

Xiang Li, Tiandi Ye, Caihua Shan, Dongsheng Li, and Ming Gao. Seegera: Self-supervised semi-implicit graph variational auto-encoders with masking. In *Proceedings of the ACM web conference 2023*, pp. 143–153, 2023.

Jungbin Lim, Jihwan Kim, Yonghyeon Lee, Cheongjae Jang, and Frank C. Park. Graph geometry-preserving autoencoders. In *Forty-first International Conference on Machine Learning, ICML 2024, Vienna, Austria, July 21-27, 2024*. OpenReview.net, 2024. URL https://openreview.net/forum?id=acTLXagzqd.

Yang Liu, Deyu Bo, Wenxuan Cao, Yuan Fang, Yawen Li, and Chuan Shi. Graph positional autoencoders as self-supervised learners. In Luiza Antonie, Jian Pei, Xiaohui Yu, Flavio Chierichetti, Hady W. Lauw, Yizhou Sun, and Srinivasan Parthasarathy (eds.), *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining, V.2, KDD 2025, Toronto ON, Canada, August 3-7, 2025*, pp. 1867–1878. ACM, 2025. doi: 10.1145/3711896.3736990. URL https://doi.org/10.1145/3711896.3736990.

Lingyu Mu, Zhengxiao Liu, Zhitong Zhu, and Zheng Lin. Trust-grs: A trustworthy training framework for graph neural network based recommender systems against shilling attacks. In Toby Walsh, Julie Shah, and Zico Kolter (eds.), *AAAI-25, Sponsored by the Association for the Advancement of Artificial Intelligence, February 25 - March 4, 2025, Philadelphia, PA, USA*, pp. 12408–12416. AAAI Press, 2025. doi: 10.1609/AAAI.V39I12.33352. URL https://doi.org/10.1609/aaai.v39i12.33352.

Thanh Tam Nguyen, Thanh Toan Nguyen, Matthias Weidlich, Jun Jo, Quoc Viet Hung Nguyen, Hongzhi Yin, and Alan Wee-Chung Liew. Handling low homophily in recommender systems with partitioned graph transformer. *IEEE Trans. Knowl. Data Eng.*, 37(1):334–350, 2025. doi: 10.1109/TKDE.2024.3485880. URL https://doi.org/10.1109/TKDE.2024.3485880.

Hongbin Pei, Bingzhe Wei, Kevin Chen-Chuan Chang, Yu Lei, and Bo Yang. Geom-gcn: Geometric graph convolutional networks. *arXiv preprint arXiv:2002.05287*, 2020.

Prithviraj Sen, Galileo Namata, Mustafa Bilgic, Lise Getoor, Brian Galligher, and Tina Eliassi-Rad. Collective classification in network data. *AI magazine*, 29(3):93–93, 2008.

Kartik Sharma, Yeon-Chang Lee, Sivagami Nambi, Aditya Salian, Shlok Shah, Sang-Wook Kim, and Srijan Kumar. A survey of graph neural networks for social recommender systems. *ACM Computing Surveys*, 56(10):1–34, 2024.

Oleksandr Shchur, Maximilian Mumme, Aleksandar Bojchevski, and Stephan Günnemann. Pitfalls of graph neural network evaluation. *arXiv preprint arXiv:1811.05868*, 2018.

Yucheng Shi, Yushun Dong, Qiaoyu Tan, Jundong Li, and Ninghao Liu. Gigamae: Generalizable graph masked autoencoder via collaborative latent space reconstruction. In *Proceedings of the 32nd ACM international conference on information and knowledge management*, pp. 2259–2269, 2024.

Lei Tang and Huan Liu. Relational learning via latent social dimensions. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 817–826, 2009.

Liang Wang, Xiang Tao, Qiang Liu, Shu Wu, and Liang Wang. Rethinking graph masked autoencoders through alignment and uniformity. In Michael J. Wooldridge, Jennifer G. Dy, and Sriraam Natarajan (eds.), *Thirty-Eighth AAAI Conference on Artificial Intelligence, AAAI 2024, Thirty-Sixth Conference on Innovative Applications of Artificial Intelligence, IAAI 2024, Fourteenth Symposium on Educational Advances in Artificial Intelligence, EAAI 2014, February 20-27, 2024, Vancouver, Canada*, pp. 15528–15536. AAAI Press, 2024. doi: 10.1609/AAAI.V38I14.29479. URL https://doi.org/10.1609/aaai.v38i14.29479.

Tingting Wang, Jiaxin Su, Haobing Liu, and Ruobing Jiang. Weighted graph structure learning with attention denoising for node classification. *arXiv preprint arXiv:2503.12157*, 2025a.

Yuhan Wang, Xiang Zhao, Zhen Tan, Weidong Xiao, and Xueqi Cheng. Temporal knowledge graph diffusion model for open-world reasoning. *Sci. China Inf. Sci.*, 68(10), 2025b. doi: 10.1007/S11432-024-4329-3. URL https://doi.org/10.1007/s11432-024-4329-3.

Haotong Wei, Yinlin Zhu, Xunkai Li, and Bin Jiang. Loyalde: Improving the performance of graph neural networks with loyal node discovery and emphasis. *Neural Networks*, 164:719–730, 2023.

Xinxu Wei, Kanhao Zhao, Yong Jiao, Hua Xie, Lifang He, and Yu Zhang. Pre-training graph contrastive masked autoencoders are strong distillers for EEG. In Aarti Singh, Maryam Fazel, Daniel Hsu, Simon Lacoste-Julien, Felix Berkenkamp, Tegan Maharaj, Kiri Wagstaff, and Jerry Zhu (eds.), *Forty-second International Conference on Machine Learning, ICML 2025, Vancouver, BC, Canada, July 13-19, 2025*, volume 267 of *Proceedings of Machine Learning Research*. PMLR / OpenReview.net, 2025a. URL https://proceedings.mlr.press/v267/wei25q.html.

Yibiao Wei, Jie Zou, Weikang Guo, Guoqing Wang, Xing Xu, and Yang Yang. MSCRS: multi-modal semantic graph prompt learning framework for conversational recommender systems. In Nicola Ferro, Maria Maistro, Gabriella Pasi, Omar Alonso, Andrew Trotman, and Suzan Verberne (eds.), *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2025, Padua, Italy, July 13-18, 2025*, pp. 42–52. ACM, 2025b. doi: 10.1145/3726302.3730040. URL https://doi.org/10.1145/3726302.3730040.

Haonan Zhang, Dongxia Wang, Zhu Sun, Yanhui Li, Youcheng Sun, Huizhi Liang, and Wenhai Wang. Kg4receval: Does knowledge graph really matter for recommender systems? *ACM Trans. Inf. Syst.*, 43(3):64:1–64:36, 2025a. doi: 10.1145/3713071. URL https://doi.org/10.1145/3713071.

Juzheng Zhang, Lanning Wei, Zhen Xu, and Quanming Yao. Heuristic learning with graph neural networks: a unified framework for link prediction. In *Proceedings of the 30th ACM SIGKDD conference on knowledge discovery and data mining*, pp. 4223–4231, 2024a.

Peiyan Zhang, Yuchen Yan, Xi Zhang, Chaozhuo Li, Senzhang Wang, Feiran Huang, and Sunghun Kim. Transgnn: Harnessing the collaborative power of transformers and graph neural networks for recommender systems. In Grace Hui Yang, Hongning Wang, Sam Han, Claudia Hauff, Guido Zuccon, and Yi Zhang (eds.), *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2024, Washington DC, USA, July 14-18, 2024*, pp. 1285–1295. ACM, 2024b. doi: 10.1145/3626772.3657721. URL https://doi.org/10.1145/3626772.3657721.

Qiqi Zhang, Chao Li, and Zhongying Zhao. Teacher-guided edge discriminator for personalized graph masked autoencoder. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 13269–13276, 2025b.

Yi Zhang, Yiwen Zhang, Dengcheng Yan, Shuiguang Deng, and Yun Yang. Revisiting graph-based recommender systems from the perspective of variational auto-encoder. *ACM Trans. Inf. Syst.*, 41(3):81:1–81:28, 2023. doi: 10.1145/3573385. URL https://doi.org/10.1145/3573385.

Yanan Zhao, Xingchao Jian, Feng Ji, Wee Peng Tay, and Antonio Ortega. Generalized graph signal reconstruction via the uncertainty principle. In *ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1–5. IEEE, 2025.

Jiajun Zhou, Shengbo Gong, Xuanze Chen, Chenxuan Xie, Shanqing Yu, Qi Xuan, and Xiaoniu Yang. Clarify confused nodes via separated learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2025.

Yanqiao Zhu, Yichen Xu, Feng Yu, Qiang Liu, Shu Wu, and Liang Wang. Graph contrastive learning with adaptive augmentation. In *Proceedings of the web conference 2021*, pp. 2069–2080, 2021.

Daniel Zügner and Stephan Günnemann. Certifiable robustness and robust training for graph convolutional networks. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, pp. 246–256, 2019.

A Proof of Proposition 1

Proof. Let $\mathbf{s}_i^{(l)} = \sum_{j\in\mathcal{N}(i)} \bar{\alpha}_{ij}^{(l)} \mathbf{W}^{(l)} \mathbf{h}_j^{(l)}$. Since perturbations are supported only on $\mathcal{V}_r$, the change in the pre-activation representation is $\Delta\mathbf{s}_i^{(l)} = \sum_{j\in\mathcal{N}(i)\cap\mathcal{V}_r} \bar{\alpha}_{ij}^{(l)} \mathbf{W}^{(l)} \Delta\mathbf{h}_j^{(l)}$. By the Lipschitz continuity of $\sigma(\cdot)$, the triangle inequality, and the induced operator norm,

$$\left\|\Delta\mathbf{h}_i^{(l+1)}\right\| \le L_\sigma \sum_{j\in\mathcal{N}(i)\cap\mathcal{V}_r} \bar{\alpha}_{ij}^{(l)} \left\|\mathbf{W}^{(l)}\right\| \left\|\Delta\mathbf{h}_j^{(l)}\right\|. \tag{21}$$

Using $\|\Delta\mathbf{h}_j^{(l)}\| \le \delta_r^{(l)}$ and the definition of $M_i^{(l)}$ yields Equation 20. □

B Datasets

To comprehensively evaluate the effectiveness and generalization capability of the proposed method, we conduct experiments on **9 representative real-world graph datasets** spanning diverse application domains, including citation networks (**Cora**, **CiteSeer**, and **PubMed** (Sen et al., 2008)), co-purchase networks (**Photo** and **Computers** (Shchur et al., 2018)), co-authorship networks (**CS** and **Physics** (Shchur et al., 2018)), and social networks (**Actor** (Pei et al., 2020) and **BlogCatalog** (Tang & Liu, 2009)). These datasets exhibit substantial diversity in node feature dimensionality, graph scale, structural sparsity, and semantic heterogeneity, enabling a rigorous evaluation of the proposed method under heterogeneous graph conditions and providing a comprehensive benchmark for assessing its robustness across different graph domains. A summary of the dataset statistics is provided in **Table 3**.

Table 3: Statistics for datasets.

Category	Datasets	Nodes	Edges	Features
Citation Networks	Cora	2708	10556	1433
	Citeseer	3327	9104	3703
	Pubmed	19717	88648	500
Co-purchase Networks	Photo	7487	119043	745
	Computers	13381	245778	767
Co-authorship Networks	CS	18333	81894	6805
	Physics	34493	247962	8415
Social Networks	Actor	7,600	33,391	931
	Blogcatalog	10,312	333,983	8,189

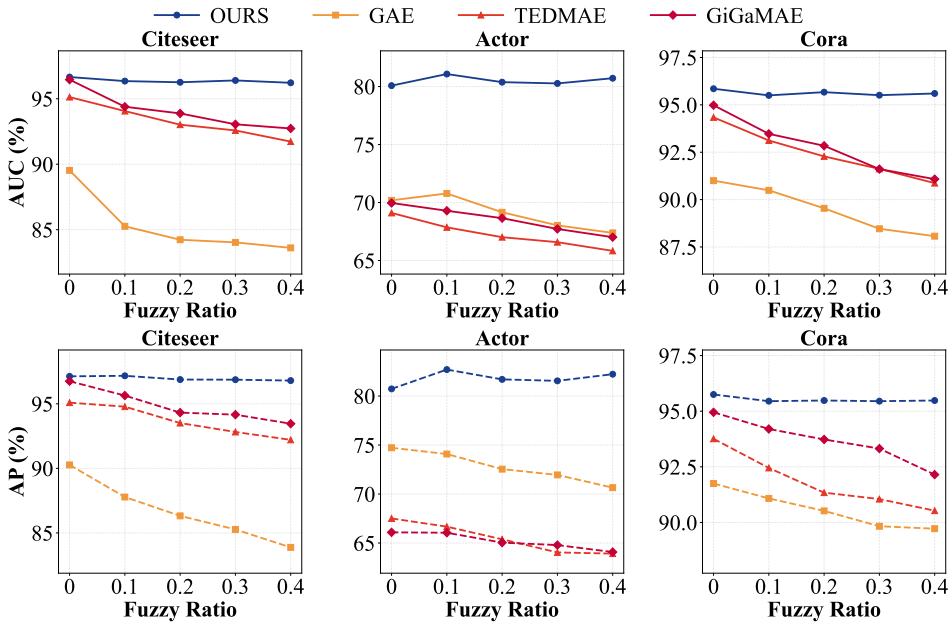


Figure 4: Link-prediction performance under controlled feature perturbations. The term "fuzzy ratio" retained in the visualization denotes the proportion of perturbed nodes rather than a ground-truth proportion of semantically unreliable nodes.

C Additional Experiments

This section provides additional analyses of SEGMA under controlled feature perturbations, structural corruption, hyperparameter variations, and different computational settings. These experiments complement the main results but should be interpreted as diagnostic stress tests. In particular, synthetic feature perturbations do not reproduce a ground-truth notion of semantic unreliability, and runtime measurements reflect the tested implementations and hardware rather than an asymptotic reduction in message-passing complexity.

C.1 Controlled Feature-Perturbation Stress Test

Experimental Protocol. We evaluate sensitivity to controlled feature perturbations on CiteSeer, Cora, and Actor. The comparison includes GAE, GiGaMAE, TED-GAE, and SEGMA. For each perturbation level, we uniformly sample a subset of nodes and add Gaussian perturbations to their input features while leaving graph topology and data splits unchanged. The perturbed-node ratio is varied over $\{0, 0.1, 0.2, 0.3, 0.4\}$, and all methods are trained from scratch under the same evaluation protocol. This experiment tests how link-prediction performance changes when an increasing proportion of input sources is perturbed; it is not intended to simulate or identify real-world semantic unreliability.

Results. As shown in Figure 4, the baseline methods generally exhibit increasing performance degradation as the perturbed-node ratio grows, whereas SEGMA shows smaller variation over the tested range. The difference is particularly visible on Actor, although slight non-monotonic changes at individual perturbation levels may also reflect optimization variability or a regularization effect introduced by feature perturbation. These results indicate that source-risk-aware message allocation is associated with greater empirical stability under this controlled setting. They do not establish robustness to arbitrary perturbations or guarantee that the same behavior holds for naturally occurring source uncertainty.

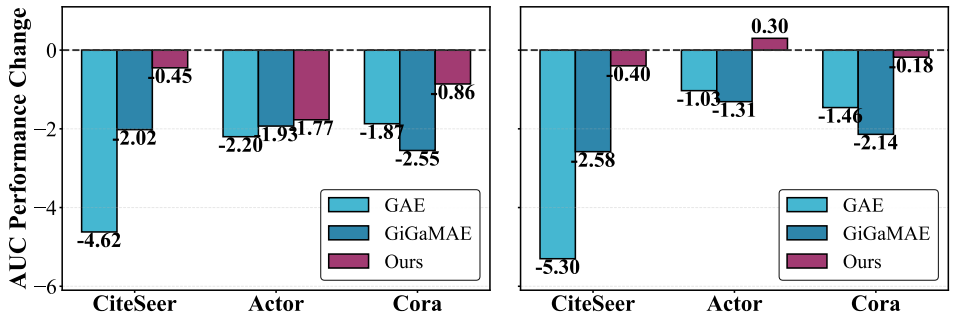


Figure 5: Sensitivity of SEGMA to structural and feature perturbations. The phrase "semantic ambiguity" retained in the visualization denotes controlled feature perturbation and should not be interpreted as ground-truth semantic ambiguity.

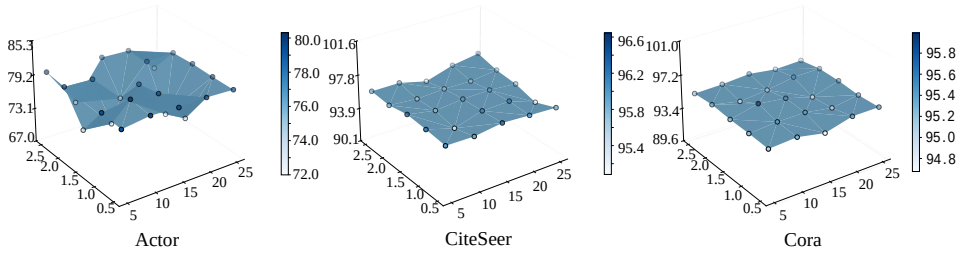


Figure 6: Sensitivity to the IQR multiplier λ and the number of frequency bands q. The parameter denoted by p in the visualization corresponds to λ in the main text.

C.2 Structural and Feature Perturbations

Experimental Protocol. We further compare the responses of SEGMA to two perturbation families. Structural corruption modifies the adjacency matrix through random edge perturbations while keeping node features fixed, whereas feature perturbation modifies selected node attributes while preserving the original topology. Experiments are conducted on CiteSeer, Cora, and Actor using the perturbation levels reported in Figure 5. The purpose is to determine whether the two perturbation types induce distinguishable empirical responses, rather than to treat either one as an exact model of source reliability.

Results. Figure 5 shows that SEGMA responds differently to structural and feature perturbations. Structural corruption produces clearer performance degradation on CiteSeer and Cora, while the response to feature perturbation is comparatively moderate over the tested range. This distinction is consistent with SEGMA preserving the observed topology and conditioning only the allocation of propagated messages. It also indicates that SEGMA should not be interpreted as a dedicated structural-denoising method. The experiment does not prove that the source-risk proxy recovers semantic uncertainty; it only shows that feature- and structure-level disturbances probe different aspects of the learned representations.

C.3 Hyperparameter Sensitivity

We study the sensitivity of SEGMA to the IQR multiplier λ and the number of frequency bands q. The former controls the selectivity of high-risk source identification, while the latter determines the resolution used in spectral aggregation. We vary $\lambda \in \{0.5, 1.0, 1.5, 2.0, 2.5\}$ and $q \in \{5, 10, 15, 20, 25\}$ on Actor, CiteSeer, and Cora, changing one parameter at a time while keeping the remaining settings fixed.

As shown in Figure 6, performance varies smoothly across the examined ranges, and no dataset exhibits an abrupt collapse under moderate parameter changes. Different datasets favor different

configurations, which is expected because their feature-score distributions and graph structures differ. Overall, the results suggest that SEGMA is not restricted to a single sharply tuned configuration within the evaluated ranges. Nevertheless, all reported test results use parameters selected according to validation performance, and this analysis should not be interpreted as evidence of parameter-free robustness.

C.4 Empirical Runtime

We report the average per-epoch training time of SEGMA, GAE, GiGaMAE, and TED-GAE on CiteSeer, Cora, and Actor under the same computing environment. These datasets have different graph sizes, feature dimensions, and edge-to-node ratios. The comparison therefore reflects practical runtime on representative graphs rather than a controlled asymptotic scaling experiment in which graph density is varied while all other factors remain fixed.

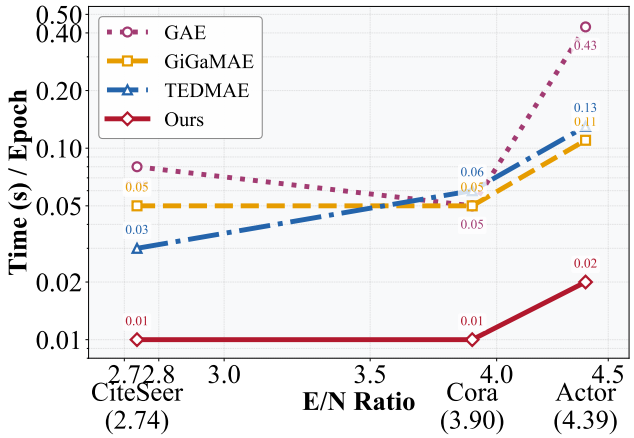


Figure 7: Average per-epoch training time on datasets with different edge-to-node ratios. Both axes follow the scaling used in the original visualization.

Figure 7 shows that SEGMA records competitive, and in these experiments lower, per-epoch training time than the compared methods. The result indicates that computing source-risk scores once before training and applying edge-level risk conditioning do not introduce prohibitive practical overhead under the tested implementation. However, SEGMA retains the original graph topology and still propagates over the observed edges; therefore, the runtime difference should not be attributed to edge pruning, fewer message-passing operations, or a lower asymptotic complexity. Differences in model architecture, implementation, preprocessing, and optimization may also contribute to the observed measurements.