Evidential Spectrum-Aware Contrastive Learning for OOD Detection in Dynamic Graphs

Nan Sun^{1,2}, Xixun Lin^{1,2} \boxtimes , Zhiheng Zhou³, Yanmin Shang^{1,2}, Zhenlin Cheng⁴ \boxtimes , and Yanan Cao^{1,2}

¹ Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China {sunnan, linxixun, shangyanmin, caoyanan}@iie.ac.cn

² School of Cyber Security, University of Chinese Academy of Sciences, Beijing, China ³ Academy of Mathematics and Systems Science, Chinese Academy of Sciences

zhouzhiheng@amss.ac.cn

⁴ Beijing Wuzi University, Beijing, China

chengzhenlin@bwu.edu.cn

Abstract. Recently, Out-of-distribution (OOD) detection in dynamic graphs, which aims to identify whether incoming data deviates from the distribution of the in-distribution (ID) training set, has garnered considerable attention in security-sensitive fields. Current OOD detection paradigms primarily focus on static graphs and confront two critical challenges: i) high bias and high variance caused by single-point estimation, which makes the predictions sensitive to randomness in the data; ii) score homogenization resulting from the lack of OOD training data, where the model only learns ID-specific patterns, resulting in overall low OOD scores and a narrow score gap between ID and OOD data. To tackle these issues, we first investigate OOD detection in dynamic graphs through the lens of Evidential Deep Learning (EDL). Specifically, we propose EviSEC, an innovative and effective OOD detector via Evidential $\underline{\mathbf{S}}$ pectrum-awar $\underline{\mathbf{E}}$ Contrastive Learning. We design an evidential neural network to redefine the output as the posterior Dirichlet distribution, explaining the randomness of inputs through the uncertainty of distribution, which is overlooked by single-point estimation. Moreover, spectrumaware augmentation module generates OOD approximations to identify patterns with high OOD scores, thereby widening the score gap between ID and OOD data and mitigating score homogenization. Extensive experiments on real-world datasets demonstrate that EviSAC effectively detects OOD samples in dynamic graphs. Our source code is available at https://github.com/Sunnan191/EviSEC.

Keywords: Dynamic graph \cdot Out-of-distribution detection \cdot Evidential deep learning \cdot Graph spectrum.

1 Introduction

Real-world graph data often evolves temporally, allowing dynamic graphs to be ubiquitously applied across non-Euclidean domains such as citation networks [1],



Fig. 1. Dynamic graph examples of (a) ID samples maintain consistent structures, while OOD samples exhibit deviations. (b) Examples include cases such as SM (Structure Manipulation) and FI (Feature Interpolation), highlighting structural changes and feature variations, respectively.

social communities [2], and transaction records [3]. Recently, studies on dynamic graphs have gained increasing popularity, among which Dynamic Graph Neural Networks (DGNNs) [4] have rapidly emerged as powerful approaches for dynamic graph representation learning [5]. These models primarily leverage Graph Neural Networks (GNNs) and sequence-based architectures to capture temporal variations in dynamic graphs.

Most existing models on dynamic graphs are predominantly trained with the closed-world assumption that the training and test data share the same data distribution, with such data termed In-Distribution (ID) data. Nevertheless, real-world scenarios frequently deviate from this ideal assumption, often involving Out-of-Distribution (OOD) dynamic graphs, which are unobserved during the training process. Fig. 1 provides examples illustrating the behavior of ID and OOD dynamic graph with two OOD types. During the inference stage, performing predictions blindly without recognizing OOD samples can render the output unreliable and pose serious risks. Such cases are especially unacceptable in safety-critical domains, such as medical diagnostics [6] and autonomous driving [7]. An ideal model should effectively handle both ID test performance and OOD detection performance. Therefore, OOD detection in dynamic graphs has seen a growing demand, which aims to identify whether incoming data deviate from the distribution of the training set.

Early works in OOD detection primarily focus on static graphs. Liu et al. [8] propose a hierarchical contrastive learning method that captures common patterns of ID graphs across different granularities (node, graph, and group levels), so that OOD graphs that deviate from these patterns can be effectively identified. GNNSAFE [9] extends energy-based models [10] to static graphs and develops an energy function that classifies nodes with high energies as OOD samples. In parallel, research on anomaly detection [11] in dynamic graphs has also gained

attraction. Netwalk [12] uses random walks and autoencoders to generate similar node representations for ID data, thereby identifying anomalous interactions between nodes with distinct representations. TADDY [13] employs transformers to capture both global and local structural anomalies in node representations. Although the above approaches perform fairly well, they are confronted with two critical limitations:

i) High bias and high variance caused by single-point estimation. Most of the aforementioned methods follow a manner of single-point estimation, which overlook the inherent randomness presented in the data, so that their OOD detection results are sensitive to natural noise, resulting in high bias and high variance in the output.

ii) Score homogenization resulting from the lack of OOD training data. During the training stage, only ID data is available whereas OOD data remains unseen. This will cause the problem of score homogenization, i.e., the model only learns the ID-specific patterns, making it tend to assign overall lower OOD scores, regardless of whether the sample is ID or OOD. The small score gap between ID and OOD data makes it challenging to effectively distinguish OOD samples.

To address these limitations, we thoroughly investigate OOD detection in dynamic graphs from a novel perspective drawing on Evidential Deep Learning (EDL)—a theoretically grounded framework for uncertainty quantification through Dirichlet distributions. Specifically, we propose EviSEC, an innovative and effective OOD detector via \underline{Evi} dential \underline{S} pectrum-awar \underline{E} Contrastive Learning. We design an evidential neural network to reshape the output class probability of single-point estimation into a probability Dirichlet distribution, allowing us to describe the variability and randomness of the data through the uncertainty of distribution. Furthermore, we develop two loss functions to guide the model to output sharper Dirichlet distributions with low uncertainty scores for ID samples while preserve the ID performance. To tackle the issue of missing OOD samples, we propose a spectral-aware data augmentation module that generates OOD approximations of dynamic graphs. Based on this, the contrastive loss can enable the model to learn and assign higher uncertainty scores, thereby mitigating the problem of score homogenization. Finally, our OOD detector computes the learned uncertainty as the OOD score, which efficiently distinguishes OOD samples with higher uncertainty. The main contributions of this study are summarized as follows:

- 1. EviSEC is the first to establish a direct link between EDL and OOD detection in dynamic graphs. We propose an evidential neural network that uses the uncertainty of posterior distribution to describe the randomness of the input, addressing the problem of single-point estimation. Our method effectively improves OOD detection while maintaining ID testing performance.
- 2. We propose a spectral-aware data augmentation module to generate OOD approximations for dynamic graphs, alleviating the lack of OOD training samples. Based on this, our contrastive learning module enlarges the score

gap, mitigating the issue of score homogenization and effectively enhancing the OOD detection performance.

3. Experimental results on multiple real-world datasets validate the claimed advantages of our approach, which achieves consistent performance gains over multiple powerful competitors with an average AUROC improvement of 24.32%.

2 Related Works

2.1 Dynamic Graph Representation Learning

Early attempts in dynamic graph representation learning [5, 14] employ traditional methods including random walks [15], matrix factorization [16], and temporal point processes [17, 18] to model graph information over time. However, compared to these methods, DGNNs outperform in expressive power, since they combine message passing and temporal modeling to better capture topological and temporal dynamics. For instance, EvolveGCN [19] employs RNNs to dynamically adapt the weights of GNNs across temporal steps. DEFT [20] utilizes transformers to model temporal dependencies through self-attention mechanisms. Instead of using RNNs, LEDG [21] applies gradient-based meta learning to learn updating strategies. While these methods excel in predictive tasks, they largely focus on ID data, often neglecting OOD detection in dynamic graphs.

2.2 Graph Out-of-distribution Detection

Graph OOD detection primarily focuses on static graphs and can be broadly categorized into the following three approaches:

Post-processing Methods. These model-agnostic methods directly process outputs of pretrained models to estimate OOD scores. For instance, Lee et al. [22] use Mahalanobis distance to measure the deviation of input data. Similarly, maximum softmax probabilities (MSP) [23] can be utilized as confidence scores, with lower values signaling OOD samples. ODIN [24] enhances MSP by incorporating temperature scaling and input perturbation to increase the score gap. Yet, these methods typically follow single-point estimation, which is prone to high variance and bias, thereby hindering their ability to recognize the OOD patterns.

Energy-based Methods. These approaches are based on the framework of energy-based models, where energy scores measure the discrepancy between input data and the model's learned distribution. Specifically, energy scores [25] effectively separate OOD samples from ID data, as OOD samples typically yield higher values. GNNSAFE [9] develops an energy function directly extracted from GNNs trained with energy propagation on static graphs. Nevertheless, directly adapting these approaches to dynamic settings is challenging since they require a fundamental redesign of the energy propagation mechanism over time.

Uncertainty-based Methods. Uncertainty [26] measures the unreliability of a model's predictive distribution. A previous study [27] adopts the graph neural stochastic diffusion framework to model uncertainty in node classification tasks. Zhao et al. [28] propose a graph-based kernel estimation method to predict nodelevel Dirichlet distributions within the EDL framework. DAEDL [29] promotes EDL performance through a density estimation algorithm. In addition, several studies [30, 31] utilize the uncertainty measurement framework of conformal prediction to provide statistical guarantees for the detection results. However, the absence of OOD samples can cause severe score homogenization in detection.

3 Preliminary

3.1 Problem Definition

The dynamic graphs are interpreted as a sequence of graphs $\mathcal{G}^{1:T} = \{\mathcal{G}^t\}_{t=1}^T$, where T specifies the total number of timesteps. Each discrete graph snapshot $\mathcal{G}^t = (\mathcal{V}^t, \mathcal{E}^t, \mathbf{X}^t, \mathbf{A}^t)$ contains node set \mathcal{V}^t , edge set \mathcal{E}^t , adjacency matrix $\mathbf{A}^t \in \{0, 1\}^{N^t \times N^t}$, and feature matrix $\mathbf{X}^t \in \mathbb{R}^{N^t \times d}$, where $N^t = |\mathcal{V}^t|$ and d respectively denote the node count and the feature dimension at timestep t. Dynamic graph representation learning aims to develop a powerful encoder $f_{\theta}(\cdot, \cdot)$ to capture the temporal dependency. To achieve this, the graph snapshots $\mathcal{G}^{t:t+\Delta t}$ are commonly employed as input to model the dynamics within the temporal window Δt . Based on such input settings, we formulate OOD detection in dynamic graphs as the task of ascertaining whether $\mathcal{G}^{t:t+\Delta t}$ follow the same distribution as ID data:

Definition 1 (OOD detection in Dynamic Graphs). Let the ID dataset \mathcal{D}_{in} (resp. OOD dataset \mathcal{D}_{out}) consist of graph sequences drawn from the distributions \mathbb{P}_{in} (resp. \mathbb{P}_{out}). The training dataset $\mathcal{D}^{tr} = \{\mathcal{G}_{in}^t\}_{t=1}^T$ is a subset of \mathcal{D}_{in} , containing ID graphs at different timesteps. The test dataset \mathcal{D}^{te} is constituted by two disjoint subsets, $\mathcal{D}_{in}^{te} \subset \mathcal{D}_{in}$ and $\mathcal{D}_{out}^{te} \subset \mathcal{D}_{out}$, i.e., $\mathcal{D}^{te} = \mathcal{D}_{in}^{te} \cup \mathcal{D}_{out}^{te}$, such that $\mathcal{D}^{tr} \cap \mathcal{D}_{in}^{te} = \emptyset$. For arbitrary input snapshots $\mathcal{G}^{t:t+\Delta t} \in \mathcal{D}^{te}$, the goal of OOD detection in dynamic graphs is to design a discriminant function $G(\cdot, \cdot)$ to determine whether $\mathcal{G}^{t:t+\Delta t}$ follow \mathbb{P}_{in} or \mathbb{P}_{out} based on the OOD detection score:

$$detect(\mathcal{G}^{t:t+\Delta t}) = \begin{cases} 1 & G\left(f_{\theta}(\mathcal{G}^{t:t+\Delta t},\mathcal{W}),\mathcal{D}^{tr}\right) \geq \gamma, \\ 0 & G\left(f_{\theta}(\mathcal{G}^{t:t+\Delta t},\mathcal{W}),\mathcal{D}^{tr}\right) < \gamma, \end{cases}$$
(1)

where γ is the detection threshold, W represents the parameters of the encoder.

3.2 Evidential Deep Learning

EDL is grounded in Subjective Logic (SL) theory [32], which represents subjective multinomial opinion as a non-negative triplet $\boldsymbol{\tau} = (\boldsymbol{b}, u, \boldsymbol{\beta})$. In a K-class classification problem, $\boldsymbol{b} = [b_1, b_2, \dots, b_K]$ assigns belief mass b_i to each class; u quantifies the overall uncertainty across classes, satisfying $u + \sum_{i=1}^{K} b_i = 1$; and

 $\boldsymbol{\beta} = [\beta_1, \beta_2, \dots, \beta_K]$ is a predefined base rate vector. Based on $\boldsymbol{\tau}$, the projected probability distribution for each class is defined as $p_i = b_i + \beta_i u$, where $i = 1, 2, \dots, K$. EDL adopts the K-dimensional Dirichlet distribution $\operatorname{Dir}^{pr}(\boldsymbol{p}; w\boldsymbol{\beta})$ as the prior distribution, where w is the prior weight. The conjugate posterior of EDL, i.e., $\operatorname{Dir}^{po}(\boldsymbol{p}; \boldsymbol{\alpha})$ with concentration (parameter) $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_K]$ is defined as:

$$\operatorname{Dir}^{po}(\boldsymbol{p};\boldsymbol{\alpha}) = \frac{\Gamma\left(\sum_{i=1}^{K} \alpha_i\right)}{\prod_{i=1}^{K} \Gamma(\alpha_i)} \prod_{i=1}^{K} p_i^{\alpha_i - 1}, \qquad (2)$$

where Γ denotes the Gamma function, $\alpha_i \geq 0$, and $p_i \neq 0$ if $\alpha_i < 1$. To obtain the posterior distribution in Eq. 2, EDL first designs a post-processing module to capture the evidence vector $\boldsymbol{e} = [e_1, e_2, \cdots, e_K]$, which represents the observation evidences for K classes. Then EDL combines the observed evidence with the prior to update the posterior concentration $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \cdots, \alpha_K]$ as

$$\alpha_i = e_i + \beta_i w, \text{ where } i = 1, 2, \dots, K.$$
(3)

Thereby the multinomial opinion $\boldsymbol{\tau} = (\boldsymbol{b}, u, \boldsymbol{\beta})$ in SL can be equivalently represented by EDL through the bijection mapping F between $\boldsymbol{\tau}$ and $\text{Dir}^{po}(\boldsymbol{p}; \boldsymbol{\alpha})$:

$$F: \boldsymbol{\tau} \longleftrightarrow \operatorname{Dir}^{po}(\boldsymbol{p}; \boldsymbol{\alpha}), \text{ where } b_i = \frac{\alpha_i - \beta_i w}{\alpha_{\operatorname{sum}}}, u = \frac{w \sum_{i=1}^K \beta_i}{\alpha_{\operatorname{sum}}}, \alpha_{\operatorname{sum}} = \sum_{i=1}^K \alpha_i.$$
(4)

EDL treats the predictive uncertainty u in Eq. 4 as a discriminative metric for OOD detection. Specifically, for ID samples shown in Fig. 2(a), the model assigns relatively high evidence to at least one class, resulting in a sharp Dirichlet distribution (large α_{sum}) with lower uncertainty ($u \rightarrow 0$). Conversely, for OOD samples shown in Fig. 2(b), the evidence allocated across all classes is relatively low, leading to a flat Dirichlet distribution (small α_{sum}) with high uncertainty ($u \rightarrow 1$). This difference in uncertainty values clearly shows how the model can tell apart ID and OOD samples. A predefined threshold γ (e.g., $\gamma = 0.5$) can thus be applied to trigger OOD detection.



Fig. 2. Heatmaps of Dirichlet distributions in 3-class classification with four groups of concentration parameters and corresponding subjective opinions. Warm (resp. cool) colors represent relatively high (resp. low) probability density values in the distribution.



Fig. 3. An overall illustration of the proposed method, EviSEC, which follows a twostage paradigm: (1) During training, original dynamic graph samples and OOD approximations generated via spectrum-aware augmentation are assigned different Dirichlet distributions by the dynamic graph encoder and the EDL module—sharper for ID samples and flatter for OOD approximations. The loss function enlarges the gap in their uncertainty scores. (2) During inference, the trained model accurately identifies OOD samples using a learned uncertainty threshold.

4 METHODOLOGY

In this section, we present a novel EDL framework for OOD detection in dynamic graphs. Our proposed model, i.e., **EviSEC** combines a flexible dynamic graph encoder (Sec. 4.1) with an evidential neural network (Sec. 4.2) to address the limitations of single-point estimation without sacrificing ID performance. Furthermore, to mitigate score homogenization, we introduce a new spectrum-aware contrastive learning strategy (Sec. 4.3). Finally, we design an OOD detector (Sec. 4.4) based on the uncertainty scores. The overall architecture is illustrated in Fig. 3.

4.1 Dynamic Graph Encoder

EviSEC adopts a dynamic graph encoder [19] to efficiently capture the topological structure and temporal dependencies in dynamic graphs through the evolution of GCN parameters. Specifically, we employ a GCN with weight matrix set $\mathcal{W}_t = \{\mathbf{W}_t^{(1)}, \mathbf{W}_t^{(2)}, \dots, \mathbf{W}_t^{(L)}\}$, where $\mathbf{W}_t^{(l)}$ denotes the weight matrix at *l*-th layer, to learn the node embedding matrix $\mathbf{Z}_t^{(l)}$ at each timestep *t*:

$$\mathbf{Z}_{t}^{(l)} = \begin{cases} \mathbf{X}^{t}, & \text{if } l = 0, \\ \sigma \left(\tilde{\mathbf{D}}^{t^{-\frac{1}{2}}} \tilde{\mathbf{A}}^{t} \tilde{\mathbf{D}}^{t^{-\frac{1}{2}}} \mathbf{Z}_{t}^{(l-1)} \mathbf{W}_{t}^{(l)} \right), & \text{if } l > 0, \end{cases}$$
(5)

where σ is a non-linear activation function, $\tilde{\mathbf{A}}^t$ is the adjacency matrix augmented with self-loops, and $\tilde{\mathbf{D}}^t$ denotes the diagonal degree matrix of $\tilde{\mathbf{A}}^t$.

To naturally store historical dependencies, $\mathbf{W}_{t}^{(l)}$ is regarded as a hidden state of the dynamical system. We utilize gated recurrent units (GRUs) to incorporate the layer input $\mathbf{Z}_{t}^{(l-1)}$ into the update of $\mathbf{W}_{t}^{(l)}$ over time:

$$\underbrace{\mathbf{W}_{t}^{(l)}}_{\text{GCN weights}} = \text{GRU}(\underbrace{\mathbf{Z}_{t}^{(l-1)}}_{\text{embeddings GCN weights}}, \underbrace{\mathbf{W}_{t-1}^{(l)}}_{\text{embeddings GCN weights}}).$$
(6)

By iteratively updating the weight matrix set from \mathcal{W}_t to $\mathcal{W}_{t+\Delta t}$, our dynamic graph encoder can aggregate the evolutionary patterns of the input $\mathcal{G}^{t:t+\Delta t}$ to generate the final output $f_{\theta}(\mathcal{G}^{t:t+\Delta t}, \mathcal{W}_{t+\Delta t}) = \mathbf{Z}_{t+\Delta t}^{(L)}$. To perform a specific classification task, the output of our encoder is passed through a classifier $g_{\phi}(\cdot)$ to generate the final prediction $\hat{y} = \operatorname{softmax} \left(g_{\phi}\left(f_{\theta}\left(\mathcal{G}^{t:t+\Delta t}, \mathcal{W}_{t+\Delta t}\right)\right)\right)$.

4.2 Evidential Neural Network

EviSEC proposes an evidential neural network to reshape the prediction $p = [p_1, p_2, \ldots, p_K]$ into a Dirichlet distribution. Thus, we can leverage the uncertainty of distribution to describe the randomness of the input, addressing the high bias and high variance issues in single-point estimation. Specifically, two loss functions are developed to guide the network to assign a sharper posterior distribution with lower uncertainty to the ID data, allowing for a robust OOD detection process. Notably, our method maintains the highest probability assigned to the target class, which preserves ID performance.

In a K-class classification task, since there is no prior information, we can simply assume the uniform Dirichlet distribution $\operatorname{Dir}^{pr}(\boldsymbol{p}; 1)$ as the prior distribution $\operatorname{Dir}^{pr}(\boldsymbol{p}; w\boldsymbol{\beta})$, where the base rate $\boldsymbol{\beta} = \begin{bmatrix} \frac{1}{K}, \frac{1}{K}, \dots, \frac{1}{K} \end{bmatrix}$ and the prior weight w = K. Based on the preliminary mentioned in Sec. 3.2, the posterior concentration $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_K]$ and the class probability $\boldsymbol{p} = [p_1, p_2, \dots, p_K]$ are derived as follows,

$$\alpha_i = e_i + \beta_i w \stackrel{\text{Dir}^{pr}(\boldsymbol{p}; \mathbf{1})}{=} e_i + 1, \text{ where } i = 1, 2, \dots, K.$$
(7)

$$p_i = \mathbb{E}_{\boldsymbol{p} \sim \text{Dir}^{po}(\boldsymbol{p}; \boldsymbol{\alpha})}[p_i] = \int p_i \text{Dir}^{po}(\boldsymbol{p}; \boldsymbol{\alpha}) \, d\boldsymbol{p} = \frac{\alpha_i}{\alpha_{\text{sum}}}, \text{where } i = 1, 2, \dots, K.$$
(8)

To update the posterior concentration $\boldsymbol{\alpha}$, our evidential neural network includes an evidence collector $\boldsymbol{e}_{\theta,\phi}(\cdot)$ to construct the corresponding evidence vector $\boldsymbol{e} = [e_1, e_2, \ldots, e_K]$:

$$\boldsymbol{e}_{\theta,\phi}\left(\mathcal{G}^{t:t+\Delta t}\right) = \exp\left(g_{\phi}\left(f_{\theta}\left(\mathcal{G}^{t:t+\Delta t},\mathcal{W}_{t+\Delta t}\right)\right)\right) - \boldsymbol{1}.$$
(9)

By subtracting all-ones vector **1** from exponentiated logits, our method ensures negative logits to yield negative evidences. To ensure that $e_{\theta,\phi}(\cdot)$ produces sharp distributions with low uncertainty scores for ID samples, our network comprises two key losses: the evidential learning cross-entropy loss \mathcal{L}_{ce-edl} and the Kullback-Leibler divergence loss \mathcal{L}_{kl} following traditional EDL framework [33]. \mathcal{L}_{ce-edl} can be calculated as the negative log-likelihood of the predicted class probability pwith respect to the one-hot label $y = [y_1, y_2, \ldots, y_K]$:

$$\mathcal{L}_{\text{ce-edl}} = \frac{1}{|\mathcal{D}^{\text{tr}}|} \sum_{(\mathcal{G}^{t:t+\Delta t}, \boldsymbol{y})\in\mathcal{D}^{\text{tr}}} \mathbb{E}_{\boldsymbol{p}\sim\text{Dir}^{po}(\boldsymbol{p};\boldsymbol{\alpha})} \left[-\sum_{i=1}^{K} y_i \log(p_i) \right]$$
$$= \frac{1}{|\mathcal{D}^{\text{tr}}|} \sum_{(\mathcal{G}^{t:t+\Delta t}, \boldsymbol{y})\in\mathcal{D}^{\text{tr}}} \sum_{i=1}^{K} y_i \mathbb{E}_{p_i\sim\text{Dir}^{po}(\boldsymbol{\alpha})} \left[-\log(p_i) \right]$$
$$= \frac{1}{|\mathcal{D}^{\text{tr}}|} \sum_{(\mathcal{G}^{t:t+\Delta t}, \boldsymbol{y})\in\mathcal{D}^{\text{tr}}} \sum_{i=1}^{K} y_i \left(\psi(\alpha_{\text{sum}}) - \psi(\alpha_i) \right).$$
(10)

The digamma function $\psi(\cdot)$ is the logarithmic derivative of the Gamma function. We can see that Eq. 10 ensures the preservation of the original ID testing performance and guides the model to assign the highest concentration to the target class. However, it may still allow $e_{\theta,\phi}(\cdot)$ to allocate relatively high evidence to non-target classes, thereby increasing the uncertainty u of ID data. To address this, we introduce \mathcal{L}_{kl} to further penalize the model if it provides excessive evidence for non-target categories:

$$\mathcal{L}_{kl} = \frac{1}{|\mathcal{D}^{tr}|} \sum_{(\mathcal{G}^{t:t+\Delta t}, \boldsymbol{y}) \in \mathcal{D}^{tr}} KL \left(\text{Dir}^{po}(\boldsymbol{p}; \hat{\boldsymbol{\alpha}}), \text{Dir}^{po}(\boldsymbol{p}; \boldsymbol{1}) \right).$$
(11)

Here $\hat{\alpha} = y + (1 - y) \odot \alpha$, and \odot represents the Hadamard product. $\hat{\alpha}$ signifies the adjusted concentration parameter, which reserves correct predictions while damping the contribution from irrelevant classes.

4.3 Spectrum-Aware Contrastive Learning

In the above computational process, since only ID data is available, the model tends to assign generally lower uncertainty scores. The absence of OOD data hinders the model's ability to learn and predict higher uncertainty scores, leading to a severe score homogenization issue and further compromising its effectiveness in distinguishing OOD data. To mitigate this, we propose a graph spectrum-aware augmentation technique to generate negative samples as OOD approximations for dynamic graphs. Through contrastive loss, we further widen the gap in uncertainty scores between ID and OOD data.

Spectrum-Aware Augmentation. Previous work [34] demonstrates that low-frequency components in the graph spectrum capture global features (e.g., graph connectivity), while high-frequency components often reflect noise. Therefore, our technique changes the low-frequency components in the graph spectrum to generate negative samples.

Let the symmetric normalized Laplacian matrix of the adjacency matrix with self-loops $\tilde{\mathbf{A}}$ represented as $\mathbf{L} = \mathbf{I} - \tilde{\mathbf{D}}^{-1/2} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-1/2}$, which can be eigendecomposed as follows,

$$\mathbf{L} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^{\top} = \sum_{i=1}^{N} \lambda_i \boldsymbol{u}_i \boldsymbol{u}_i^{\top}, \qquad (12)$$

where $\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_N)$ is the eigenvalue diagonal matrix and $\mathbf{U} = [\mathbf{u}_1^{\top}, \mathbf{u}_2^{\top}, \dots, \mathbf{u}_N^{\top}]$ is the corresponding orthogonal eigenvector matrix. For simplicity's sake, we assume $0 \leq \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N < 2$. Then the graph spectrum can be partitioned into two parts: low-frequency components with eigenvalues $\boldsymbol{\lambda}_{low} = \{\lambda_1, \lambda_2, \dots, \lambda_{\lfloor N/2 \rfloor}\}$ and high-frequency components with eigenvalues $\boldsymbol{\lambda}_{high} = \{\lambda_{\lfloor N/2 \rfloor+1}, \lambda_{\lfloor N/2 \rfloor+2}, \dots, \lambda_N\}.$

Our method generates OOD approximations by perturbing the low-frequency information with a preservation ratio r, where $0 \le r < 1$. For the low-frequency eigenspaces $\{\boldsymbol{u}_i \boldsymbol{u}_i^{\mathsf{T}}\}_{i=1}^{\lfloor N/2 \rfloor}$ of $\boldsymbol{\lambda}_{low}$, we only preserve the first r portion and discard the remaining part. For the high-frequency eigenspaces $\{\boldsymbol{u}_j \boldsymbol{u}_j^{\mathsf{T}}\}_{j=\lfloor N/2 \rfloor+1}^N$ of $\boldsymbol{\lambda}_{high}$, we keep them intact. Based on this, the negative sample \mathbf{L}_r^- is generated as follows,

$$\mathbf{L}_{r}^{-} = \sum_{i=1}^{\lfloor rN/2 \rfloor} \boldsymbol{u}_{i} \boldsymbol{u}_{i}^{\top} + \sum_{j=\lfloor N/2 \rfloor+1}^{N} \boldsymbol{u}_{j} \boldsymbol{u}_{j}^{\top}.$$
 (13)

Contrastive Loss. With the above negative samples, we seek to enable the model to learn and assign higher uncertainty scores for OOD approximations, which correspond to smoother posterior distributions. To increase the gap in uncertainty scores between the sample pairs, our contrastive loss is formed as the log-likelihood between the class probability $\mathbf{p}^- = [p_1^-, p_2^-, \dots, p_K^-]$ of the negative sample and the class probability \mathbf{p} of the original input:

$$\mathcal{L}_{\rm cl} = \frac{1}{|\mathcal{D}^{\rm tr}|} \sum_{(\mathcal{G}^{t:t+\Delta t}, \boldsymbol{y})\in\mathcal{D}^{\rm tr}} \mathbb{E}_{\boldsymbol{p}^{-}} \left[\sum_{i=1}^{K} p_i \log(p_i^{-}) \right].$$
(14)

With two balancing factors ρ_1 and ρ_2 , the overall loss function of EviSEC is formulated as:

$$\mathcal{L} = \mathcal{L}_{ce-edl} + \rho_1 \mathcal{L}_{kl} + \rho_2 \mathcal{L}_{cl}.$$
 (15)

4.4 OOD Detector

After our model has been optimized by Eq. 15, EviSEC includes an OOD detector to compute the uncertainty of posterior distribution $\text{Dir}^{po}(\boldsymbol{p}; \boldsymbol{\alpha})$ as the OOD score. In specific, the uncertainty u is derived from the Dirichlet concentration $\boldsymbol{\alpha}$ through Eq. 4:

$$u = \frac{K}{\alpha_{\text{sum}}}, \text{ where } \alpha_{\text{sum}} = \sum_{i=1}^{K} \alpha_i.$$
 (16)

On the one hand, a lower u corresponds to a sharper and more confident output distribution, suggesting the sample is from the ID class. On the other hand, a higher u corresponds to a smoother output distribution, indicating that the sample is likely an OOD instance. Then the OOD detector can select a threshold γ to determine whether the input is OOD according to the learned u.

5 EXPERIMENT

This section empirically validates the effectiveness of EviSEC by addressing four key research questions: **RQ1**, how does EviSEC perform compared with competitive methods? **RQ2**, how do the evidential neural network and the spectrumaware contrastive learning enhance OOD detection in dynamic graphs? **RQ3**, how effective is our method in staying compatible with ID tasks? **RQ4**, how do key parameters affect EviSEC's performance?

Table 1. Dataset details. "Time Splits" shows the data division (train/val/test).

	$\#\mathbf{Nodes}$	$\# \mathbf{Edges}$	$\# Time \ Splits$	Task
BC-OTC [3]	5,881	35,588	$95\ /\ 14\ /\ 28$	Edge Classification
BC-Alpha [3]	3,777	24,173	$95\ /\ 13\ /\ 28$	Edge Classification
UCI [2]	$1,\!899$	59,835	$62 \;/\; 9 \;/\; 17$	Link Prediction
AS [35]	$6,\!474$	$13,\!895$	$70\ /\ 10\ /\ 20$	Link Prediction
Elliptic [36]	203,769	$234,\!355$	$31\ /\ 5\ /\ 13$	Node Classification
Brain [37]	5,000	$1,\!955,\!488$	$10 \ / \ 1 \ / \ 1$	Node Classification

5.1 Experimental Setup

Datasets. As summarized in Tab. 1, we adopt the datasets and preprocessing steps used in previous dynamic graph representation works [19–21]. These include: i) **BC-OTC**, a trust-based bitcoin transaction network; ii) **BC-Alpha**, a similar bitcoin transaction network from another platform; iii) **UCI**, a student community network; iv) **AS**, a router traffic flow network; v) **Elliptic**, bitcoin transactions from the Elliptic network; and vi) **Brain**, a region connectivity network in the brain.

OOD Data. We extend the OOD data generation framework from recent work [9] on static graphs to generate OOD dynamic graphs through two strategies: i) Structure Manipulation (**SM**), which utilizes a stochastic block model to generate graphs as OOD samples, and ii) Feature Interpolation (**FI**), which applies random interpolation to create node features for OOD data.

Baselines. Beyond the three OOD detection families outlined in our Related Works, we also utilize several dynamic graph anomaly detection methods as baselines. These include: i) post-processing methods (MSP [23], ODIN [24], Mahalanobis [22]), ii) energy-based methods (Energy [25], GNNSAFE [9]), iii) uncertainty-based methods (Entropy [38], EDL [39], DAEDL [29]) and iv) anomaly detection methods(NetWalk [12], TADDY [13], SLADE [40]).

Metrics. We adopt the commonly used evaluation metrics [41] for OOD detection, including: i) **AUROC**, the area under the receiver operating characteristic curve; ii) **AUPR**, area under the precision-recall curve; and iii) **FPR95**, false positive rate at 95% true positive rate. Additionally, **F1**, the harmonic mean of precision and recall, is employed to evaluate ID performance.

Implementation. We employ grid search to determine the optimal values of three key hyper-parameters in our model: the balancing factors ρ_1 and ρ_2 , and the preservation rate r in augmentation. The remaining parameters, including the training epochs, the number of layers L, hidden dimension and temporal window size Δt , are adopted from prior studies [19, 20]. To ensure a fair comparison, both the baselines and EviSEC are configured with identical settings. All experiments were conducted on Python 3.6.13 and NVIDIA A800 80GB GPU.

5.2 OOD Detection Perfrmance(RQ1)

In Tab. 2 and Tab. 3, we present the key results of EviSEC in comparison with eleven competitive models from four families for the SM and FI OOD types, respectively. These experiments are conducted across six real-world scenarios, measured by AUROC, AUPR, and FPR95. For the OOD type of SM, EviSEC consistently surpasses all baselines across all six datasets, boosting the average AUROC by 24.57%, increasing the average AUPR by 25.64%, and reducing the average FPR95 by 34.65%. For the OOD type of FI, the corresponding changes are 24.08%, 25.22%, and 37.39%, respectively. These results demonstrate that our model exhibits strong adaptability to different OOD types. In both OOD types, we are pleased to observe exceptionally satisfactory performance on the BTC-OTC and BC-Alpha datasets. Specifically, we achieved nearly perfect AUROC and AUPR scores, as well as extremely low FPR95, significantly surpassing the performance of other models. These results reinforce the superiority of EviSEC for OOD detection in dynamic graphs.

5.3 Ablation Study(RQ2)

To systematically evaluate the contribution of core components (\mathcal{L}_{ce-edl} , \mathcal{L}_{kl} and \mathcal{L}_{cl}) in our framework, we conducted comprehensive ablation studies across six benchmark datasets. Through progressive removal of individual loss components from the complete EviSEC architecture, we quantitatively assessed their respective impacts on OOD detection capability measured by **AUROC**. As presented

Table 2. OOD detection results for the OOD type of structure manipulation (SM) evaluated by AUROC/AUPR/FPR95 on six datasets. The best performance is **bolded**, with the runner-up <u>underlined</u>.

Dataset	BC-OT	С І	BC-Alpha	UC	I	AS		Elliptic	Br	ain
Baseline	•			AUROC	↑ AUPI	R↑ FPF	195 ↓			
NetWalk TADDY SLADE	52.38 48.06 77.20 65.38 74.35 64.30	92.80 48.62 98.07 64.28 94.26 66.40	52.70 98.40 59.05 95.20 60.50 92.80	48.67 46. 54.80 58. 59.80 50.	$\begin{array}{c ccccc} 02 & 98.68 & 5 \\ 69 & 95.18 & 5 \\ 20 & 94.19 & 5 \end{array}$	2.71 56.31 6.89 55.73 1.20 56.98	94.20 38.9 88.67 52.6 89.20 48.8	7 46.58 91.30 1 54.98 93.20 0 54.70 95.20	54.38 52 56.80 59 51.30 55	.60 100.0 .08 100.0 .47 <u>95.58</u>
MSP ODIN Mahalanobis	$\left \begin{array}{rrrr} 48.65 & 52.82 \\ 55.15 & 42.63 \\ 61.44 & 34.88 \end{array}\right $	92.55 45.79 88.88 55.26 96.80 43.21	48.63 91.50 49.07 100.0 39.34 92.34	54.38 46. 41.68 39. 40.68 38.	$\begin{array}{c cccc} 01 & 97.20 & 5 \\ 02 & 96.68 & 5 \\ 72 & 94.29 & 5 \end{array}$	2.76 47.39 6.08 45.25 7.26 49.77	94.20 48.6 98.34 48.7 97.04 49.5	4 40.43 90.25 7 50.62 92.57 6 52.03 98.62	44.87 42 48.83 42 36.20 32	.15 99.80 .39 100.0 .17 98.65
Energy GNNSAFE	58.80 31.88 87.02 85.57	88.11 50.33 85.32 <u>89.66</u>	52.68 99.19 86.34 78.20	50.68 55. 62.41 <u>66.</u>	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	7.21 54.23 7.08 59.71	$\begin{array}{c c c c c c c c c c c c c c c c c c c $	9 52.70 98.37 8 60.22 96.55	$\begin{vmatrix} 34.67 & 48 \\ 72.12 & \underline{71} \end{vmatrix}$.94 99.27 .17 98.87
Entropy EDL DAEDL	$\begin{array}{ c c c c c c c c c c c c c c c c c c c$	99.09 50.55 98.40 55.33 69.88 85.34	55.18 100.0 54.08 100.0 80.20 90.68	60.27 50. 55.18 47. 70.22 64.	66 99.21 5 02 96.55 5 09 90.27 7	5.39 60.22 8.77 62.72 4.48 69.85	92.44 52.6 90.58 54.8 88.36 <u>66.5</u>	6 55.71 98.06 7 50.68 97.54 3 68.97 80.65	$\begin{array}{ c c c c c c c c c c c c c c c c c c c$.33 99.80 .60 99.40 .17 96.31
EviSEC	96.39 94.52	13.47 97.14	92.30 13.86	72.09 68.	84 82.77 7	8.39 76.50	80.46 70.6	1 72.35 74.68	76.73 73	$.51\ 92.64$

Table 3. OOD detection results for the OOD type of feature interpolation (**FI**) evaluated by AUROC/AUPR/FPR95 on six datasets. The best performance is **bolded**, with the runner-up <u>underlined</u>.

Dataset	BC-	-OTC	E	C-Alph	ıa		UCI			AS			Elliptic			Brain	
Baseline	·					AUI	$ROC\uparrow$	AU	$PR\uparrow$	FPF	195 ↓						
NetWalk	57.06 5	8.54 91.78	60.45	65.12	99.20	50.86	48.34	90.60	56.27	50.20	88.71	42.08	45.32	99.80	42.48	37.08	100.0
TADDY	79.68 6	4.43 94.07	68.31	66.61 58.06	96.08	56.76	52.63 49.06	94.71	55.16	51.20	85.64	50.65	55.08 58.62	96.20	52.99 47.03	55.03	100.0
SLADE	1 10.01 0	4.09 93.20	05.80	38.00	94.02	04.02	49.00	90.55	30.34	55.98	90.70	32.74	36.02	94.04	47.03	55.95	98.40
MSP	43.04 3	8.80 94.38	3 46.28	38.98	94.34	33.74	42.58	98.77	51.20	48.29	99.25	44.28	39.38	88.64	34.18	33.66	100.0
ODIN	50.26 4	7.87 90.32	2 54.20	50.84	99.80	46.20	48.07	98.82	58.91	47.30	96.53	50.24	59.26	94.40	50.07	56.30	99.25
Mahalanobis	45.20 5	1.35 98.20	45.84	42.09	95.63	43.82	44.24	93.29	56.98	51.28	95.77	48.53	54.08	89.60	44.31	42.07	98.65
Energy	60.28 5	9.60 72.37	51.20	59.61	95.44	55.21	56.38	90.52	54.30	52.27	94.70	58.89	60.58	94.20	48.67	50.38	97.27
GNNSAFE	70.05 <u>8</u>	<u>6.79</u> 68.30	87.25	88.37	68.46	64.09	68.27	89.42	66.57	64.35	80.29	66.80	58.92	94.38	63.44	60.18	<u>90.43</u>
Entropy	42.95 2	6.83 100.0	47.23	50.35	100.0	55.69	54.31	98.70	52.17	48.69	88.64	59.07	62.18	96.50	70.02	71.17	96.56
EDL	53.81 3	8.03 95.42	2 59.30	52.18	100.0	48.67	49.72	94.21	64.38	64.80	82.50	61.20	64.52	90.20	64.20	68.17	93.57
DAEDL	80.54 8	6.28 <u>54.3</u>	80.14	74.58	65.31	65.28	64.97	91.25	70.07	$\underline{69.57}$	86.25	68.34	66.30	78.57	72.18	73.15	96.51
EviSEC	97.29 9	8.51 5.22	96.40	94.44	10.27	70.91	72.08	80.16	72.58	74.25	84.20	72.15	70.25	58.68	74.21	76.47	89.40

in Tab. 4, the full configuration demonstrates significant superiority over all ablated variants, with particular note to the Elliptic dataset where we observe a substantial absolute performance improvement of 17.41%. This empirical validation confirms that the synergistic combination significantly enhances OOD detection performance.

5.4 ID Performance(RQ3)

As mentioned in Sec. 4.2, EviSEC does not alter the value of the maximum component in the predictive distribution. Therefore, when comparing ID performance measured by **F1** scores with representation learning methods such as EvolveGCN [19], LEDG [21], DEFT-MLP, DEFT-GAT and DEFT-T [20], our approach still achieves performance close to (or even surpassing) state-of-the-art levels (see Fig. 4). This demonstrates that our method does not sacrifice ID task performance while enhancing OOD detection capabilities.

$\mathcal{L}_{\mathrm{ce-edl}}$	$\mathcal{L}_{\mathrm{kl}}$	$\mathcal{L}_{\mathrm{cl}}$	BC-OTC	BC-Alpha	UCI	AS	Elliptic	Brain
-	\checkmark	\checkmark	86.24	85.14	66.28	65.80	58.32	68.45
\checkmark	-	\checkmark	95.58	95.28	70.60	72.87	68.24	73.12
\checkmark	\checkmark	-	90.17	93.69	68.07	74.10	65.78	69.56
\checkmark	\checkmark	×	91.08	92.26	69.31	73.54	67.92	70.27
\checkmark	\checkmark	\checkmark	96.39	97.14	72.09	78.39	70.61	76.73
	0.8 - <u>-</u> 0.6 - □ 0.4 - 0.2 -	1					Evolv LEOG DEFF DEFF DEFF Evise	egCN MLP GAT T C
	0.2	BC-	отс	BC-Alpha	Ellipt	ic	Brain	

Table 4. Ablation study results for the SM OOD type across six datasets. \checkmark means the variant with random negative edge sampling augmentation.

Fig. 4. In-Distribution performance of classification tasks with F1 scores on the y-axis.

5.5 Parameter Sensitivity(RQ4)

Fig. 5 shows the hyper-parameter sensitivity of EviSEC measured by AUROC in terms of the balancing factors ρ_1 and ρ_2 , and the preservation rate r in augmentation. Specifically, ρ_1, ρ_2 range from 0.2 to 2.0 with a step size of 0.2 on the AS dataset, and $r \in \{0, 0.2, 0.4, 0.6, 0.8\}$ on six datasets. We observe that setting them to moderate values (e.g., $\rho_1 \in \{0.4, 0.6, 0.8\}, \rho_2 \in \{0.6, 0.8, 1.0\}, r \in$ [0.2, 0.4]) usually achieves optimal performance. We can conclude that: i) \mathcal{L}_{ce-edl} and \mathcal{L}_{cl} have a greater impact than \mathcal{L}_{kl} ; and ii) a smaller r, i.e., retaining less lowfrequency information, tends to generate more effective OOD approximations.



Fig. 5. Hyper-parameter sensitivity of EviSEC.

6 Conclusion

In this work, we propose EviSEC, an innovative and effective OOD detection framework comprising two key modules. Specifically, we propose an evidential neural network that uses the uncertainty of posterior distribution to explain the randomness of the input, thereby mitigating the high bias and high variance issues associated with single-point estimation. Moreover, we design a spectralaware contrastive learning module to generate OOD approximations for enlarging the score gap between the ID and OOD data, effectively mitigating the issue of score homogenization. Empirical evaluation demonstrates that our model outperforms others across various datasets for OOD detection in dynamic graphs.

Acknowledgments. This work is supported by the National Key Research and Development Program of China (NO.2022YFB3102200) and the National Natural Science Foundation of China (No.62402491).

References

- Tang, J., Zhang, J., Yao, L., Li, J., Zhang, L., Su, Z.: Arnetminer: extraction and mining of academic social networks. In: Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining. pp. 990–998 (2008)
- 2. Rossi, R.A., Ahmed, N.K.: The network data repository with interactive graph analytics and visualization. In: AAAI (2015), https://networkrepository.com
- Kumar, S., Hooi, B., Makhija, D., Kumar, M., Faloutsos, C., Subrahmanian, V.: Rev2: Fraudulent user prediction in rating platforms. In: Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining. pp. 333–341 (2018)
- Zhang, S., Chen, L., Wang, C., Li, S., Xiong, H.: Temporal graph contrastive learning for sequential recommendation. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 38, pp. 9359–9367 (2024)
- Khoshraftar, S., An, A.: A survey on graph representation learning methods. ACM Transactions on Intelligent Systems and Technology 15(1), 1–55 (2024)
- Kononenko, I.: Machine learning for medical diagnosis: history, state of the art and perspective. Artificial Intelligence in medicine 23(1), 89–109 (2001)
- Yurtsever, E., Lambert, J., Carballo, A., Takeda, K.: A survey of autonomous driving: Common practices and emerging technologies. IEEE access 8, 58443–58469 (2020)
- Liu, Y., Ding, K., Liu, H., Pan, S.: Good-d: On unsupervised graph out-ofdistribution detection. In: Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining. pp. 339–347 (2023)
- 9. Wu, Q., Chen, Y., Yang, C., Yan, J.: Energy-based out-of-distribution detection for graph neural networks. arXiv preprint arXiv:2302.02914 (2023)
- Ranzato, M., Boureau, Y.L., Chopra, S., LeCun, Y.: A unified energy-based framework for unsupervised learning. In: Artificial Intelligence and Statistics. pp. 371– 379. PMLR (2007)
- Song, C., Lin, X., Shen, H., Shang, Y., Cao, Y.: Uniform: Towards unified framework for anomaly detection on graphs. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 39, pp. 12559–12567 (2025)

- 16 Nan Sun et al.
- Yu, W., Cheng, W., Aggarwal, C.C., Zhang, K., Chen, H., Wang, W.: Netwalk: A flexible deep embedding approach for anomaly detection in dynamic networks. In: Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining. pp. 2672–2681 (2018)
- Liu, Y., Pan, S., Wang, Y.G., Xiong, F., Wang, L., Chen, Q., Lee, V.C.: Anomaly detection in dynamic graphs via transformer. IEEE Transactions on Knowledge and Data Engineering 35(12), 12081–12094 (2021)
- Lin, X., Cao, J., Zhang, P., Zhou, C., Li, Z., Wu, J., Wang, B.: Disentangled deep multivariate hawkes process for learning event sequences. In: 2021 IEEE International Conference on Data Mining (ICDM). pp. 360–369. IEEE (2021)
- Grover, A., Leskovec, J.: node2vec: Scalable feature learning for networks. In: Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining. pp. 855–864 (2016)
- Li, J., Dani, H., Hu, X., Tang, J., Chang, Y., Liu, H.: Attributed network embedding for learning in a dynamic environment. In: Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. pp. 387–396 (2017)
- Zuo, Y., Liu, G., Lin, H., Guo, J., Hu, X., Wu, J.: Embedding temporal network via neighborhood formation. In: Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining. pp. 2857–2866 (2018)
- 18. Zhang, S., Zhou, C., Liu, Y., Zhang, P., Lin, X., Ma, Z.M.: Neural jump-diffusion temporal point processes. In: International Conference on Machine Learning (2024)
- Pareja, A., Domeniconi, G., Chen, J., Ma, T., Suzumura, T., Kanezashi, H., Kaler, T., Schardl, T., Leiserson, C.: Evolvegcn: Evolving graph convolutional networks for dynamic graphs. In: Proceedings of the AAAI conference on artificial intelligence. vol. 34, pp. 5363–5370 (2020)
- Bastos, A., Nadgeri, A., Singh, K., Suzumura, T., Singh, M.: Learnable spectral wavelets on dynamic graphs to capture global interactions. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 37, pp. 6779–6787 (2023)
- Xiang, X., Huang, T., Wang, D.: Learning to evolve on dynamic graphs (student abstract). In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 36, pp. 13091–13092 (2022)
- Lee, K., Lee, K., Lee, H., Shin, J.: A simple unified framework for detecting outof-distribution samples and adversarial attacks. Advances in neural information processing systems **31** (2018)
- Hendrycks, D., Gimpel, K.: A baseline for detecting misclassified and out-ofdistribution examples in neural networks. arXiv preprint arXiv:1610.02136 (2016)
- Liang, S., Li, Y., Srikant, R.: Enhancing the reliability of out-of-distribution image detection in neural networks. arXiv preprint arXiv:1706.02690 (2017)
- Liu, W., Wang, X., Owens, J., Li, Y.: Energy-based out-of-distribution detection. Advances in neural information processing systems 33, 21464–21475 (2020)
- Abdar, M., Pourpanah, F., Hussain, S., Rezazadegan, D., Liu, L., Ghavamzadeh, M., Fieguth, P., Cao, X., Khosravi, A., Acharya, U.R., et al.: A review of uncertainty quantification in deep learning: Techniques, applications and challenges. Information fusion **76**, 243–297 (2021)
- Lin, X., Zhang, W., Shi, F., Zhou, C., Zou, L., Zhao, X., Yin, D., Pan, S., Cao, Y.: Graph neural stochastic diffusion for estimating uncertainty in node classification. In: Forty-first International Conference on Machine Learning (2024)
- Zhao, X., Chen, F., Hu, S., Cho, J.H.: Uncertainty aware semi-supervised learning on graph data. Advances in neural information processing systems 33, 12827–12836 (2020)

- Yoon, T., Kim, H.: Uncertainty estimation by density aware evidential deep learning. In: International Conference on Machine Learning. pp. 57217–57243. PMLR (2024)
- Zhang, S., Zhou, C., Liu, Y., Zhang, P., Lin, X., Pan, S.: Conformal anomaly detection in event sequences. In: International Conference on Machine Learning (2025)
- Lin, X., Cao, Y., Sun, N., Zou, L., Zhou, C., Zhang, P., Zhang, S., Zhang, G., Wu, J.: Conformal graph-level out-of-distribution detection with adaptive data augmentation. In: Proceedings of the ACM on Web Conference 2025. pp. 4755– 4765 (2025)
- 32. Jøsang, A.: Subjective logic, vol. 3. Springer (2016)
- Sensoy, M., Kaplan, L., Kandemir, M.: Evidential deep learning to quantify classification uncertainty. Advances in neural information processing systems 31 (2018)
- Liu, N., Wang, X., Bo, D., Shi, C., Pei, J.: Revisiting graph contrastive learning from the perspective of graph spectrum. Advances in Neural Information Processing Systems 35, 2972–2983 (2022)
- Leskovec, J., Kleinberg, J., Faloutsos, C.: Graphs over time: densification laws, shrinking diameters and possible explanations. In: Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining. pp. 177–187 (2005)
- Weber, M., Domeniconi, G., Chen, J., Weidele, D.K.I., Bellei, C., Robinson, T., Leiserson, C.E.: Anti-money laundering in bitcoin: Experimenting with graph convolutional networks for financial forensics. arXiv preprint arXiv:1908.02591 (2019)
- 37. Xu, D., Cheng, W., Luo, D., Gu, Y., Liu, X., Ni, J., Zong, B., Chen, H., Zhang, X.: Adaptive neural network for node classification in dynamic networks. In: 2019 IEEE International Conference on Data Mining (ICDM). pp. 1402–1407. IEEE (2019)
- Shannon, C.E.: Communication theory of secrecy systems. The Bell system technical journal 28(4), 656–715 (1949)
- Deng, D., Chen, G., Yu, Y., Liu, F., Heng, P.A.: Uncertainty estimation by fisher information-based evidential deep learning. In: International Conference on Machine Learning. pp. 7596–7616. PMLR (2023)
- 40. Lee, J., Kim, S., Shin, K.: Slade: Detecting dynamic anomalies in edge streams without labels via self-supervised learning. In: Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. pp. 1506–1517 (2024)
- Guo, Y., Yang, C., Chen, Y., Liu, J., Shi, C., Du, J.: A data-centric framework to endow graph neural networks with out-of-distribution detection ability. In: Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. pp. 638–648 (2023)