doi:10.3772/j.issn.1002 - 0470.2026.06.003

基于区间边界传播的 DeepWalk 可验证鲁棒性①

俞山青② 方 虚 王金焕 彭松涛

(浙江工业大学信息工程学院 杭州 310023)

摘 要 基于 DeepWalk 的图嵌入算法的非线性和复杂性使其可验证鲁棒性的研究变得十分困难。因此,本文提出了一种基于区间边界传播(interval bound propagation,IBP)的方法,在节点度值攻击下验证 DeepWalk 在链路预测任务上的鲁棒性。本文将 DeepWalk 等效于对目标矩阵的奇异值分解任务,根据 IBP 的思想,首先将原始度值矩阵的扰动范围传递到目标矩阵,再传递到点的嵌入向量,获取其扰动区间作为约束条件,最终通过求解优化问题得到鲁棒验证结果。最后,利用本文提出的方法,对不同图网络规模、图网络密度和嵌入向量维数对可验证鲁棒节点对占比的影响进行了实验。结果显示平均鲁棒验证率最高可达 63.63%。

关键词 DeepWalk;链路预测;区间边界传播法;可验证鲁棒性

近年来,随着图数据的广泛应用,图嵌入技术成为研究领域中备受关注的焦点之一。在图嵌入方法中,DeepWalk 以其在学习节点表示方面的卓越表现而备受瞩目[1]。同时,DeepWalk 被证明其本质上可以看作是对目标矩阵进行矩阵分解[2]。这种表示不仅能够捕捉节点之间的关系,还能为下游任务提供有力的特征支持,例如链路预测[3]。

链路预测涉及揭示缺失的链接或预测节点之间的交互,在网络科学中具有广泛的应用,如网络重构[4]、评估网络演化机制[5]、推荐系统[6-7]等。尽管 DeepWalk 在链路预测上取得了显著的成功,但在实际应用中,图数据可能会受到各种威胁和攻击,例如节点、连边的删除或增加等,这可能导致 DeepWalk 学到的表示失去准确性和鲁棒性[8],对下游任务造成较大影响。针对 DeepWalk 的鲁棒性的提升,目前已有相关研究。例如 Dai 等人[9]提出了一种基于对抗网络的嵌入算法框架,它利用对抗学习原则来增强 DeepWalk 等嵌入算法的鲁棒性。该框架由 2 个部分组成,即保持结构的部分和对抗学习的部分。前者旨在捕捉网络结构属性,而后者旨在将潜在表示的后验分布与给定的先验相匹配。Dai 等人[10]采用对抗性训练的思想,通过在训练过程中引入对抗样本,迫使模型更好地捕捉数据的特征,从而提高 DeepWalk 在真实场景和面对攻击时的鲁棒性。然而,这些属于基于经验的启发式的方式往往容易受到更强大的攻击而被扰动成功,而鲁棒性验证这种方式可以终结这种可能性。

为得到 DeepWalk 的可验证鲁棒性结果,本文基于区间边界传播思想提出了一种针对矩阵分解形式的 DeepWalk 鲁棒性验证方法。本文首先分析节点度值的扰动范围,然后将这个扰动传播至 DeepWalk 的目标矩阵,随后将其进一步传播至奇异值分解(singular value decomposition,SVD)后的奇异值和奇异向量中,最终影响到每个节点的嵌入向量。其中针对 SVD 这个非线性环节,本文引用了一种适用于未知但有界(unknown-but-bounded,UBB)参数的 SVD 区间扰动方法[11],用以确定扰动边界。通过计算对每个节点嵌入向量的扰动边界,得到一个具有

① 国家自然科学基金(62103374,U21B2001)和浙江省重点研发计划(2024C01025,2022C01018)资助项目。
② 女,1984 年生,博士,副教授;研究方向:智能计算,数据挖掘;联系人,E-mail:yushanqing@zjut.edu.cn。
(收稿日期:2024 - 03 - 29)

约束条件的优化问题,从而完成对 DeepWalk 模型鲁棒性的验证。综上所述,本文的主要贡献如下。

(1)提出了一种基于区间边界传播思想的针对 DeepWalk 矩阵分解形式的鲁棒性验证方法,该方法完成了 DeepWalk 模型在面对节点度值攻击时的鲁棒性验证。

(2)在扰动传播过程中,利用 DeepWalk 的等价矩阵分解形式,避免了随机游走等随机方式的复杂性;针对 SVD 这一非线性环节,引用了适用于未知但有界参数的 SVD 区间扰动方法,用以确定扰动边界。

(3)针对不同网络规模、网络密度和嵌入向量维数对可验证鲁棒节点对占比的影响进行了实验。

1 相关研究

目前针对图对抗样本的鲁棒性验证问题主要分为 2 种方法:区间边界传播法(interval bound propagation,IBP)与随机平滑法(randomized smoothness,RS)。

IBP 是一种用于推导输入和输出变量之间关系的算法。它通过计算约束条件的上下界来推导变量的可能取值范围。Sunaga[12] 和 Gowal 等人[13] 通过引入基于 IBP 思想的可验证鲁棒神经网络训练方法,以确保在输入数据受到扰动时,网络仍能在 ∞ 范数限制内保持可靠性。Zügner 和 Günnemann[14-15] 将这种思想应用到图网络领域,以解决图卷积神经网络(graph convolutional networks,GCN)模型的可验证鲁棒性问题。他们利用 IBP 思想来推导输入样本的可扰动范围,并将其传播到 GCN 模型中。为解决非线性部分传播问题,研究者采用了凸松弛技术[16],并将整个问题转化为最优化问题以进行量化求解。

虽然 IBP 方法具有较快的速度,但其对模型特性的高度依赖使得对于不同模型需要具体分析和选择适当的计算方式,尤其是在处理非线性环节时可能会相当复杂。

RS 的基本思想是通过在输入数据上引入随机扰动,将原始模型转变为随机模型,并且具有可验证的鲁棒性。Lecuyer 等人[17] 提出了 PixelDP(pixel differential privacy),利用差分隐私和模型鲁棒性之间的联系首次提出了随机平滑的方法,提供了对 top-1 预测进行随机平滑处理的鲁棒性验证。Jia 等人[18] 和 Wang 等人[19] 将 RS 技术成功应用于图网络领域的可验证鲁棒性中,对图神经网络在处理图结构数据的节点和图分类任务以及社区发现任务上的可验证鲁棒性问题进行了论证,使得原本在连续数据上使用的 RS 成功地应用于图结构这种离散数据,并取得了显著的效果。

相较于 IBP,RS 方法不用依赖于模型的结构而直接地去研究它的鲁棒性,不仅更加方便,也能适用于更大规模的任务。然而,这种方法实质上验证的并不是原始模型而是加入噪声后的随机模型。

综上所述,为了对 DeepWalk 原始模型的鲁棒性验证问题进行研究,本文将从区间边界传播的角度出发建立有效的鲁棒性验证方法。

2 DeepWalk 的鲁棒性验证方法

为应对前文描述的困难与挑战,本节设计了 DeepWalk 的鲁棒性验证方法的具体流程。该方法框架主要分为 3 个部分:鲁棒性验证对应的优化问题建模,链路预测任务的设定,以及扰动边界的传播。首先,明确定义了与鲁棒性验证问题相关的优化问题,其中的目标函数是基于链路预测任务的预测函数构建的,而约束条件则是从度值矩阵扰动传播到节点嵌入向量的结果。整体框架如图 1 所示。

2.1 优化问题定义

为了得到鲁棒性验证结果,本文首先需要建立一个鲁棒性验证对应的优化问题,再通过求解这个优化问题得到最后的鲁棒性验证结果。

设一个无向网络由 $G=(V,E)$ 表示,其中 V 和 E 分别表示节点和连边的集合,假定节点编号的集合为 $V=\{1,\cdots,N\}$。通过 $\Omega=\{(n_i,n_j)\mid n_i,n_j\in V, n_i\neq n_j\}$ 定义网络中所有节点对的集合。在无向网络中,连边 (n_i,n_j) 与连边 (n_j,n_i) 相同,并且 Ω 包含 $\binom{N}{2}=\frac{|V|\cdot(|V|-1)}{2}$ 个元素。

假设输入的图结构 G 已确定,它的邻接矩阵为 $\boldsymbol{A}\in\{0,1\}^{N\times N}$,相应的度值矩阵为 $\boldsymbol{D}=\mathrm{diag}(d_1,d_2,\cdots,d_N)$,其中 d_i 是节点 i 的度值,$d_i=\sum_{j=1}^{N}A_{ij}$。在这

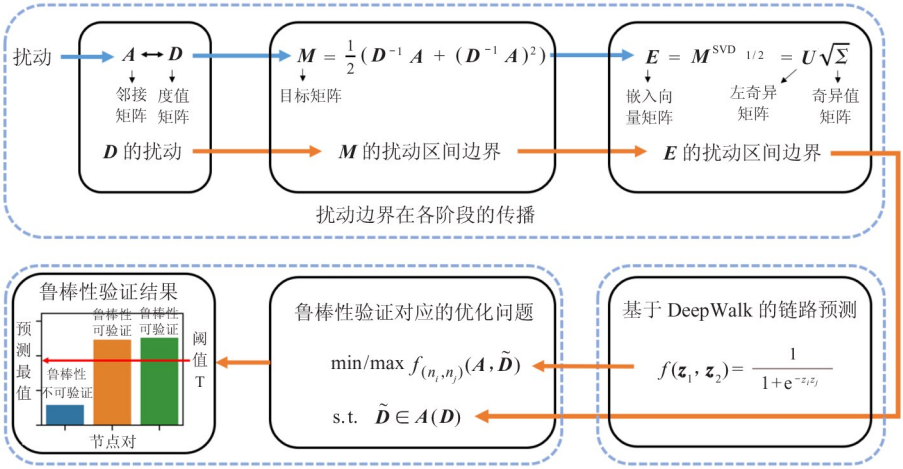

图 1　DeepWalk 鲁棒性验证方法框架

里，规定扰动只能对某个目标节点 t 的任意一个一跳邻居节点的度值增加 ϵ，ϵ 表示度值增加量。这样被扰动后的度值矩阵 $\tilde{\boldsymbol{D}}$ 未知。本文的目标是验证图中每个节点对的鲁棒性。这样的证明确保对节点对 (n_i,n_j) 的预测不能被攻击者改变。表 1 为本文所使用的符号描述。以下是任意节点对鲁棒性验证对应的优化问题表达式。

表 1　符号描述

符号	描述
$\tilde{\boldsymbol{D}}$	被攻击后的度值矩阵
$\mathscr{A}(\boldsymbol{D})$	度值矩阵在所有可能攻击后组成的矩阵集合
$\lvert V\rvert$	图网络中所有连边的个数
$\lvert\Omega\rvert$	图网络中所有节点可以组成的节点对的个数
N_r	可验证鲁棒节点对数量
$\boldsymbol{M}^{\mathrm{I}}$	扰动传递后的目标矩阵，矩阵中每个元素为一个取值区间
$\boldsymbol{M}^{\mathrm{c}}$	$\boldsymbol{M}^{\mathrm{I}}$ 中每个元素的中间值
$\Delta\boldsymbol{M}$	$\boldsymbol{M}^{\mathrm{I}}$ 中每个元素的半径范数

问题 1　给定图 G，其邻接矩阵 $\boldsymbol{A}$ 和目标节点 t。在度值矩阵 $\boldsymbol{D}$ 被扰动后产生的扰动集 $\mathscr{A}(\boldsymbol{D})$ 下，有如下情况。

（1）在未被扰动情况下，当节点对 (n_i,n_j) 的链路预测结果为存在连边，即 $f_{(n_i,n_j)}(\boldsymbol{A},\boldsymbol{D}) \geqslant T$ 时，链路预测值的最小值为

$$\min f_{(n_i,n_j)}(\boldsymbol{A},\tilde{\boldsymbol{D}}) \tag{1}$$

$$\text{s.t.}\quad \tilde{\boldsymbol{D}} \in \mathscr{A}(\boldsymbol{D}) \tag{2}$$

式中：$f_{(n_i,n_j)}(\,)$ 为计算节点 n_i、n_j 之间是否存在连边的预测值的函数，T 为链路预测任务中判断是否存在连边的阈值，具体计算方法如后文式（10）、（11）所示。当 $\min f_{(n_i,n_j)}(\boldsymbol{A},\tilde{\boldsymbol{D}}) \geqslant T$，则节点对 (n_i,n_j) 在扰动集 $\mathscr{A}(\boldsymbol{D})$ 下是鲁棒的。

（2）在未被扰动情况下，当节点对 (n_i,n_j) 的链路预测结果为不存在连边，即 $f_{(n_i,n_j)}(\boldsymbol{A},\boldsymbol{D}) < T$ 时，链路预测值的最大值为：

$$\max f_{(n_i,n_j)}(\boldsymbol{A},\tilde{\boldsymbol{D}}) \tag{3}$$

$$\text{s.t.}\quad \tilde{\boldsymbol{D}} \in \mathscr{A}(\boldsymbol{D}) \tag{4}$$

当 $\max f_{(n_i,n_j)}(\boldsymbol{A},\tilde{\boldsymbol{D}}) < T$，则节点对 (n_i,n_j) 在扰动集 $\mathscr{A}(\boldsymbol{D})$ 下是鲁棒的。

2.2　基于矩阵分解的 DeepWalk 的链路预测

基于负采样的 skip-gram（skip-gram with negative sampling，SGNS）是一种用于学习自然语言处理中词嵌入的算法。DeepWalk 算法对其进行修改，使其从语义网络泛化到各种普通图网络。随后，Levy 和 Goldberg[20] 证明了 SGNS 词向量表示学习算法实质上是矩阵分解一个偏移正点对互信息（shifted postive pointwise mutual information，SPPMI）矩阵，简称为 $\boldsymbol{B}$ 矩阵。

$$B_{ij} = \log\frac{N(n_i,c_j) \times \lvert W\rvert}{N(n_i) \times N(c_j)} - \log s \tag{5}$$

式中：B_{ij} 代表 $\boldsymbol{B}$ 矩阵中第 i 行、第 i 列的元素，s 代表每个词语对的负采样个数，W 表示整个训练集中词语的数量，$N(n_i)$ 表示词语 n_i 在整个训练集 W 中的出现次数，$N(c_j)$ 表示上下文词语 c_j 在整个训练

集 W 中的出现次数, $N(n_i,c_j)$ 表示词语对在整个训练集中的出现次数。

随后,Yang 和 Liu[21] 证明了 DeepWalk 算法相当于分解 SGNS 的特征矩阵。

$$F_{ij} = \log \frac{N(n_i,c_j)}{N(n_i)} \tag{6}$$

式中: F_{ij} 代表 $\boldsymbol{F}$ 矩阵中第 i 行、第 j 列的元素, n_i 表示图结构中的节点,不再是词语, c_j 表示上下文节点。

最后,Yang 等人[22] 权衡了算法速度与准确度两方面,得到需要分解的目标矩阵为式(7)。

$$\boldsymbol{M} = \frac{1}{2}(\boldsymbol{D}^{-1}\boldsymbol{A} + (\boldsymbol{D}^{-1}\boldsymbol{A})^2) \tag{7}$$

式中: $\boldsymbol{A}$ 为图的邻接矩阵, $\boldsymbol{D}$ 为其对应的度值矩阵。

矩阵分解方式可采用 SVD。相较于一些涉及概率模型或复杂正则化的其他分解方法,SVD 在数学上更加直观和可解释,适用于需要清晰数学基础的应用场景。此外,SVD 适用于多种类型的矩阵,包括稠密和稀疏矩阵,并且无需额外的前提假设或调整。目标矩阵 $\boldsymbol{M}$ 经过 SVD 分解后可得式(8)。

$$\boldsymbol{M} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}} \tag{8}$$

式中:上标 T 表示矩阵的转置。

最终的嵌入向量矩阵可表示为式(9)。

$$\boldsymbol{E} = \boldsymbol{U}\sqrt{\boldsymbol{\Sigma}} \tag{9}$$

式中: $\boldsymbol{U}$ 为对 $\boldsymbol{M}$ 进行 SVD 分解得到的左奇异矩阵, $\sqrt{\boldsymbol{\Sigma}}$ 指对奇异值矩阵 $\boldsymbol{\Sigma}$ 中所有的值进行开方运算。 $\boldsymbol{E}$ 为嵌入向量矩阵,矩阵的每行对应相应节点的嵌入向量。对于大规模矩阵,SVD 的计算复杂度较高,其时间复杂度为 $O(N^3)$。

得到节点嵌入向量后,即可进行链路预测任务。本文使用 sigmoid 函数来计算 2 个节点的相似度以预测它们之间是否存在连边。当节点 n_i、n_j 的嵌入向量分别为 $\boldsymbol{z}_i$、$\boldsymbol{z}_j$, 它们之间存在连边的预测值为

$$f(\boldsymbol{z}_i,\boldsymbol{z}_j) = \frac{1}{1+\mathrm{e}^{-\boldsymbol{z}_i\boldsymbol{z}_j}} \tag{10}$$

当 $f(\boldsymbol{z}_i,\boldsymbol{z}_j) \geqslant T$ 时,模型预测节点对之间存在连边;当 $f(\boldsymbol{z}_i,\boldsymbol{z}_j) < T$ 时,则它们之间不存在连边。所有节点对的预测值组成的数列记为 P。

阈值 T 的计算方法为:对所有 $|\Omega|$ 个节点对的链路预测的预测值数列 P 进行降序排序得到 P_{sorted}, 其中第 $|V|$ 个值记为阈值 T, 如式(11)所示。

$$T = P_{\text{sorted}}[|V|] \tag{11}$$

2.3 扰动边界传播

本文将扰动边界的传播过程分成 2 个部分:度值矩阵传播到 DeepWalk 的目标矩阵($\boldsymbol{D}\rightarrow\boldsymbol{M}$)以及目标矩阵传播到节点嵌入向量矩阵($\boldsymbol{M}\rightarrow\boldsymbol{E}$)。

2.3.1 $\boldsymbol{D}\rightarrow\boldsymbol{M}$ 阶段

将需要进行矩阵分解的目标矩阵 $\boldsymbol{M}$ 展开后,可以得到矩阵中每个元素 M_{ij} 的表达式如式(12)所示。

$$M_{ij} = \begin{cases} \dfrac{1}{2}\left(\dfrac{A_{i1}}{d_i d_1} + \dfrac{A_{i2}}{d_i d_2} + \cdots + \dfrac{A_{ik}}{d_i d_k}\right) & i = j \\ \dfrac{1}{2}\left(\dfrac{A_{ij}}{d_i} + \dfrac{A_{i1}A_{1j}}{d_i d_1} + \cdots + \dfrac{A_{ik}A_{kj}}{d_i d_k}\right) & i \neq j \end{cases} \tag{12}$$

式中: $k \in \{1,2,\cdots,N\}$ 且 $k \neq j$。当攻击方式为对某个目标节点的某个一跳邻居节点的度值扰动时,对 M_{ij} 的扰动需要进行分类讨论。具体分为以下 3 种情况。

(1)对于目标节点 t 对应的行中的元素,由于扰动仅涉及目标节点的一跳邻居节点,因此目标节点的度值 d_i 保持不变,能改变的只可能为 d_k 中的对应于目标节点的一跳邻居节点的度值。为了最大化元素的变化,需要考虑在目标节点的一跳邻居节点中选择哪一个节点来增加其度值。因为 $g(x) = \dfrac{1}{x}$, $x > 0, x \in \mathbb{N}$ 是一个单调递减函数,且其二阶导数为正,说明当 x 增加时,函数值的下降逐渐减缓,即 $\dfrac{1}{u} - \dfrac{1}{u+1} > \dfrac{1}{u+1} - \dfrac{1}{u+2}(u \in \mathbb{N} > 1)$, 所以当 x 越小时,增加 x 的值对 $\dfrac{1}{x}$ 的影响越大。因此,需要找到目标节点的一个邻居节点中度值最小的节点,对其增加 ϵ, 即可得到该行中 $\tilde{\boldsymbol{M}}_{ij}$ 的扰动上界和下界,如式(13)所示。

$$\begin{cases} \dfrac{1}{2}\left(\dfrac{A_{i1}}{d_i d_1} + \cdots + \dfrac{A_{im}}{d_i(d_m+\epsilon)} + \cdots + \dfrac{A_{ik}}{d_i d_k}\right) \\ \qquad \leqslant \tilde{\boldsymbol{M}}_{ij} \leqslant M_{ij} \quad i = j \\ \dfrac{1}{2}\left(\dfrac{A_{ij}}{d_i} + \dfrac{A_{i1}A_{1j}}{d_i d_1} + \cdots + \dfrac{A_{im}A_{mj}}{d_i(d_m+\epsilon)} + \cdots + \dfrac{A_{ik}A_{kj}}{d_i d_k}\right) \\ \qquad \leqslant \tilde{\boldsymbol{M}}_{ij} \leqslant M_{ij} \quad i \neq j \end{cases} \tag{13}$$

式中：m 为目标节点的一跳邻居节点中度值最小的节点。

(2)另一部分可能受到扰动的行为目标节点的一跳邻居节点所对应的行。在这种情况下，能改变的为 d_i 或者 d_k，而这些行中的每个元素都包含有 d_i，显然增加它就可以得到最大的扰动。此时，这些行中 $\tilde{\boldsymbol{M}}_{ij}$ 的扰动的上界和下界如式(14)所示。

$$\begin{cases}\dfrac{1}{2}\left(\dfrac{A_{i1}}{(d_i+\epsilon)d_1}+\dfrac{A_{i2}}{(d_i+\epsilon)d_2}+\cdots+\dfrac{A_{ik}}{(d_i+\epsilon)d_k}\right) \\ \qquad\qquad \leqslant \tilde{\boldsymbol{M}}_{ij} \leqslant M_{ij} \quad i=j \\ \dfrac{1}{2}\left(\dfrac{A_{ij}}{d_i+\epsilon}+\dfrac{A_{i1}A_{1j}}{(d_i+\epsilon)d_1}+\cdots+\dfrac{A_{ik}A_{kj}}{(d_i+\epsilon)d_k}\right) \\ \qquad\qquad \leqslant \tilde{\boldsymbol{M}}_{ij} \leqslant M_{ij} \quad i\neq j\end{cases} \tag{14}$$

(3)最后，其他可能受到扰动的行是目标节点的二跳邻居节点(不包括一跳邻居节点)所对应的行。此时，仍然只需要考虑 d_k，那么 $\tilde{\boldsymbol{M}}_{ij}$ 的扰动的上界和下界与式(13)相同。

2.3.2　$\boldsymbol{M}\rightarrow\boldsymbol{E}$ 阶段

本文中 DeepWalk 的矩阵分解形式采用 SVD。然而，SVD 扰动的不确定性量化计算是很困难的。为解决这个问题，本文引用了一种 Yang 和 Shi[11] 提出的适用于 UBB 参数的 SVD 的区间扰动方法。与需要确定参数统计数据的概率方法不同，这种方法只需要确定参数的上下界。通过使用非概率理论，这种方法可以提供奇异值和奇异向量的准确且快速的计算。下面简单介绍一下这种方法。

SVD 的原始定义如下：对于一个确定的实矩阵 $\boldsymbol{X}\in\mathbb{R}^{i\times j}$，($i$，$j$ 分别表示矩阵的行数与列数)有 $\boldsymbol{X}=\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}}$，其中，$\boldsymbol{\Sigma}$ 为奇异值矩阵，$\boldsymbol{U}$、$\boldsymbol{V}$ 分别是左、右奇异矩阵。组成它们的元素都是确定的实数。但是对于一个不确定性矩阵，其中每个元素都可以被看作是一个 UBB 参数。上一小节所求得的 $\tilde{\boldsymbol{M}}_{ij}$ 的扰动上界和下界刚好可以构成这样的不确定性矩阵 $\boldsymbol{M}^{\mathrm{I}}$。由于矩阵由有界不确定性定义，奇异值分解结果也由 UBB 参数表征。这样的矩阵可以表示为确定性部分和不确定性部分的组合：

$$\boldsymbol{M}^{\mathrm{I}}=\boldsymbol{M}^{\mathrm{c}}+\Delta\boldsymbol{M}^{\mathrm{I}}=\boldsymbol{M}^{\mathrm{c}}+[-1,1]\Delta\boldsymbol{M}=[\underline{\boldsymbol{M}},\overline{\boldsymbol{M}}] \tag{15}$$

$$M_{pq}^{\mathrm{I}}=M_{pq}^{\mathrm{c}}+\Delta M_{pq}^{\mathrm{I}}=M_{pq}^{\mathrm{c}}+[-1,1]\Delta M_{pq}=[\underline{M_{pq}},\overline{M_{pq}}] \tag{16}$$

式中：M_{pq}^{I} 是不确定区间 $\boldsymbol{M}^{\mathrm{I}}$ 中的第 p 行第 q 列元素。元素 M_{pq}^{c} 和矩阵 $\boldsymbol{M}^{\mathrm{c}}$ 是确定性部分，在区间数学理论中也称为中心值和中心矩阵。$\Delta\boldsymbol{M}^{\mathrm{I}}$ 是不确定部分，$\Delta\boldsymbol{M}$ 是其边界，也称为半径范数。ΔM_{pq} 是 $\Delta\boldsymbol{M}$ 中的第 p 行第 q 列元素。$\underline{\boldsymbol{M}}$ 和 $\overline{\boldsymbol{M}}$ 是 $\boldsymbol{M}^{\mathrm{I}}$ 的下界和上界。$\underline{M_{pq}}$ 和 $\overline{M_{pq}}$ 是 M_{pq}^{I} 的下界和上界。该方法的目标是解决具有 UBB 参数的奇异值分解，表示为：

$$\boldsymbol{M}^{\mathrm{I}}=\boldsymbol{U}^{\mathrm{I}}\boldsymbol{\Sigma}^{\mathrm{I}}(\boldsymbol{V}^{\mathrm{I}})^{\mathrm{T}} \tag{17}$$

式中：$\boldsymbol{\Sigma}^{\mathrm{I}}=\begin{bmatrix}\boldsymbol{\Sigma}_r^{\mathrm{I}} & 0\\ 0 & 0\end{bmatrix}$，$\boldsymbol{\Sigma}_r^{\mathrm{I}}=\mathrm{diag}(\sigma_1^{\mathrm{I}},\sigma_2^{\mathrm{I}},\cdots,\sigma_r^{\mathrm{I}})$。确定部分为 $\boldsymbol{U}^{\mathrm{c}}$，$\boldsymbol{u}_i^{\mathrm{c}}$，$\boldsymbol{V}^{\mathrm{c}}$，$\boldsymbol{v}_i^{\mathrm{c}}$，$\boldsymbol{\Sigma}^{\mathrm{c}}$，$\boldsymbol{\Sigma}_r^{\mathrm{c}}$ 和 $\{\sigma_1^{\mathrm{I}},\sigma_2^{\mathrm{I}},\cdots,\sigma_r^{\mathrm{I}}\}$；不确定部分为 $\Delta\boldsymbol{U}^{\mathrm{I}}$，$\Delta\boldsymbol{u}_i^{\mathrm{I}}$，$\Delta\boldsymbol{V}^{\mathrm{I}}$，$\Delta\boldsymbol{v}_i^{\mathrm{I}}$，$\Delta\boldsymbol{\Sigma}^{\mathrm{I}}$，$\Delta\boldsymbol{\Sigma}_r^{\mathrm{I}}$ 和 $\{\Delta\sigma_1^{\mathrm{I}},\Delta\sigma_2^{\mathrm{I}},\cdots,\Delta\sigma_r^{\mathrm{I}}\}$；以及不确定区间为 $\boldsymbol{U}^{\mathrm{I}}$，$\boldsymbol{u}_i^{\mathrm{I}}$，$\boldsymbol{V}^{\mathrm{I}}$，$\boldsymbol{v}_i^{\mathrm{I}}$，$\boldsymbol{\Sigma}^{\mathrm{I}}$，$\boldsymbol{\Sigma}_r^{\mathrm{I}}$ 和 $\{\sigma_1^{\mathrm{I}},\sigma_2^{\mathrm{I}},\cdots,\sigma_r^{\mathrm{I}}\}$；它们的半径为 $\Delta\boldsymbol{U}$，$\Delta\boldsymbol{u}_i$，$\Delta\boldsymbol{V}$，$\Delta\boldsymbol{v}_i$，$\Delta\boldsymbol{\Sigma}$，$\Delta\boldsymbol{\Sigma}_r$ 和 $\{\Delta\sigma_1,\Delta\sigma_2,\cdots,\Delta\sigma_r\}$。这些确定部分、不确定部分之间以及半径的关系可以通过与式(15)、式(16)相同的方式得到。

当存在扰动 $\Delta\boldsymbol{M}^{\mathrm{I}}$ 时，奇异值分解的奇异值和向量可以通过扰动方法求解，通过保留一阶范数并忽略高阶范数，利用区间扩展，可以得到奇异值的半径范数：

$$\Delta\sigma_i=|(u_i^c)^{\mathrm{T}}|\sum_{p=1}^{m}\sum_{q=1}^{n}\left|\frac{\partial M(M_{pq}^{\mathrm{c}})}{\partial M_{pq}}\right|\Delta M_{pq}|\boldsymbol{v}_i^{\mathrm{c}}| \tag{18}$$

奇异值的下界和上界可由式(19)得到。

$$\begin{cases}\underline{\sigma_i}=\sigma_i^{\mathrm{c}}-\Delta\sigma_i\\ \overline{\sigma_i}=\sigma_i^{\mathrm{c}}+\Delta\sigma_i\end{cases} \tag{19}$$

对于奇异向量，按照确定奇异值的上、下界的方法，可以获得奇异向量的半径范数为

$$\Delta u_i=\sum_{p=1}^{m}\sum_{q=1}^{n}\left|\frac{\partial u_i(M^{\mathrm{c}})}{\partial M_{pq}}\right|\Delta M_{pq} \tag{20}$$

奇异向量的下界和上界可由式(21)得到。

$$\begin{cases}\underline{u_i}=u_i^{\mathrm{c}}-\Delta u_i\\ \overline{u_i}=u_i^{\mathrm{c}}+\Delta u_i\end{cases} \tag{21}$$

在得到奇异值和奇异向量的上、下界后，根据

式(9)可得到图网络中每个节点的嵌入向量的上、下界。这样,每个节点对 n_i、n_j 的扰动后的嵌入向量 $\tilde{z}_i$、$\tilde{z}_j$ 的上、下界将作为问题 1 的约束条件。同时,结合式(10)作为目标函数,可将问题 1 转化为

$$\min/\max \quad \frac{1}{1+\mathrm{e}^{-\tilde{z}_i\tilde{z}_j}} \tag{22}$$

$$\text{s.t.} \quad \underline{z_i} \leqslant \tilde{z}_i \leqslant \overline{z_i}, \underline{z_j} \leqslant \tilde{z}_j \leqslant \overline{z_j} \tag{23}$$

上述问题为简单的有界线性规划问题,容易求解。本文算法的伪代码总结如算法 1 所示。

算法 1 基于区间边界传播的鲁棒性验证

输入:图网络 G,目标节点 t,扰动预算 ϵ

输出:可验证鲁棒节点对占比 R

根据式(7)~式(9),计算原始嵌入向量矩阵 $\boldsymbol{E}$
根据式(10),计算阈值 T
根据式(12)~式(21),计算扰动后嵌入向量的上下界 $\underline{z_i} \leqslant \tilde{z}_i \leqslant \overline{z_i}$
可验证鲁棒节点对数 N_r 置 0
for 每对节点对 $(n_i, n_j) \in \Omega$ **do**
    if $\frac{1}{1+\mathrm{e}^{-z_iz_j}} \geqslant T$ **then**
        计算 $\min \frac{1}{1+\mathrm{e}^{-\tilde{z}_i\tilde{z}_j}}$, s.t. $\underline{z_i} \leqslant \tilde{z}_i \leqslant \overline{z_i}, \underline{z_j} \leqslant \tilde{z}_j \leqslant \overline{z_j}$
        if $\min \frac{1}{1+\mathrm{e}^{-\tilde{z}_i\tilde{z}_j}} \geqslant T$ **then**
            N_r += 1
        end if
    else
        计算 $\max \frac{1}{1+\mathrm{e}^{-\tilde{z}_i\tilde{z}_j}}$, s.t. $\underline{z_i} \leqslant \tilde{z}_i \leqslant \overline{z_i}, \underline{z_j} \leqslant \tilde{z}_j \leqslant \overline{z_j}$
        if $\max \frac{1}{1+\mathrm{e}^{-\tilde{z}_i\tilde{z}_j}} < T$ **then**
            N_r += 1
        end if
    end if
end for
return $R = \frac{N_r}{|\Omega|}$

综上所述,本文将在度值攻击下验证 DeepWalk 的鲁棒性的问题表述为一个优化问题,并通过链路预测的预测函数和扰动边界传播得到优化问题的目标函数和约束条件,形成一个高效的鲁棒性验证过程。接下来,本文将通过一系列实验,研究在不同条件下 DeepWalk 的鲁棒性验证结果。

3 实验结果与分析

本节将通过实验研究在不同图网络规模、密度以及嵌入向量维度下,DeepWalk 的鲁棒性验证结果的变化趋势。

3.1 实验数据以及评价指标

本实验的实验平台是 NVIDIA Tesla V100-PCIE,算法程序运行于 Linux 系统,Python3.9 版本。在具体实验中,本文采用随机方式生成图网络数据。随机生成边的方式使得可以在一定范围内灵活地控制节点数量和边数,从而生成不同规模和结构的图网络。这样可以得到多样性的网络样本,有助于对不同情况下的网络行为和性能进行全面的分析和比较。具体地,为确保图网络的全连通性以维持网络的完整性,实验中采用了环形图作为起始结构,以保证图中的所有节点均为连通状态,避免出现孤立节点。随后,通过随机生成边的方式创建了多个图网络,节点数量分别为 50、60、70、80,边数分别为 60、70、80、90、100、110、120、130、140 和 150。每个节点数量和边数的组合生成了 20 个实例,实验结果取这些实例的均值,以确保有足够的样本进行可靠的分析。注意在网络生成过程中,应避免自环和重复的边。目标节点的编号被设定为 20。通过这样的实验设计,旨在研究不同规模和密度的图网络对于鲁棒性验证结果的影响,为实验结果提供更为可靠的分析基础。

实验采用可验证鲁棒节点对占比 R 作为评价指标,当可验证鲁棒的节点对的数量为 N_{r},则可验证鲁棒节点对占比为

$$R = \frac{N_{\mathrm{r}}}{|\Omega|} \tag{24}$$

式中:$|\Omega|$ 为图网络中所有节点对的数量。该指标用于量化模型在鲁棒性验证中验证成功的节点对数量相对于整个节点对集合的比例。

3.2 实验结果对比

现有的针对图网络领域的可验证鲁棒性研究方法主要分为 2 大类:区间边界传播法和随机平滑法。

随机平滑法是一种可以应用于图网络领域的技术,其核心思想是在输入数据上引入一定程度的随机扰动,从而将原始模型转变为随机模型,并且具有

可验证鲁棒性的特点。这种可验证的鲁棒性是通过对随机扰动的范围和影响进行严格的数学分析和验证来实现的。

图 2 展示了本文提出的方法与随机平滑方法的实验结果。从图 2 可以看出:(1)随着扰动大小的增加,2 种方法的鲁棒验证率都呈现下降趋势。其中,随机平滑法的下降幅度更为剧烈。(2)在各种规模的图网络中,在扰动预算为 1 时,随机平滑法的效果远好于区间边界传播法,在每种图网络中都取得了超过 80% 的鲁棒验证率,尤其在平均度较大的图网络中效果更为显著。(3)随着扰动大小增加,区间边界传播法优于随机平滑法,尤其在平均度较低的图网络中更加显著。这证明了本文方法在面对较大扰动时的优势。

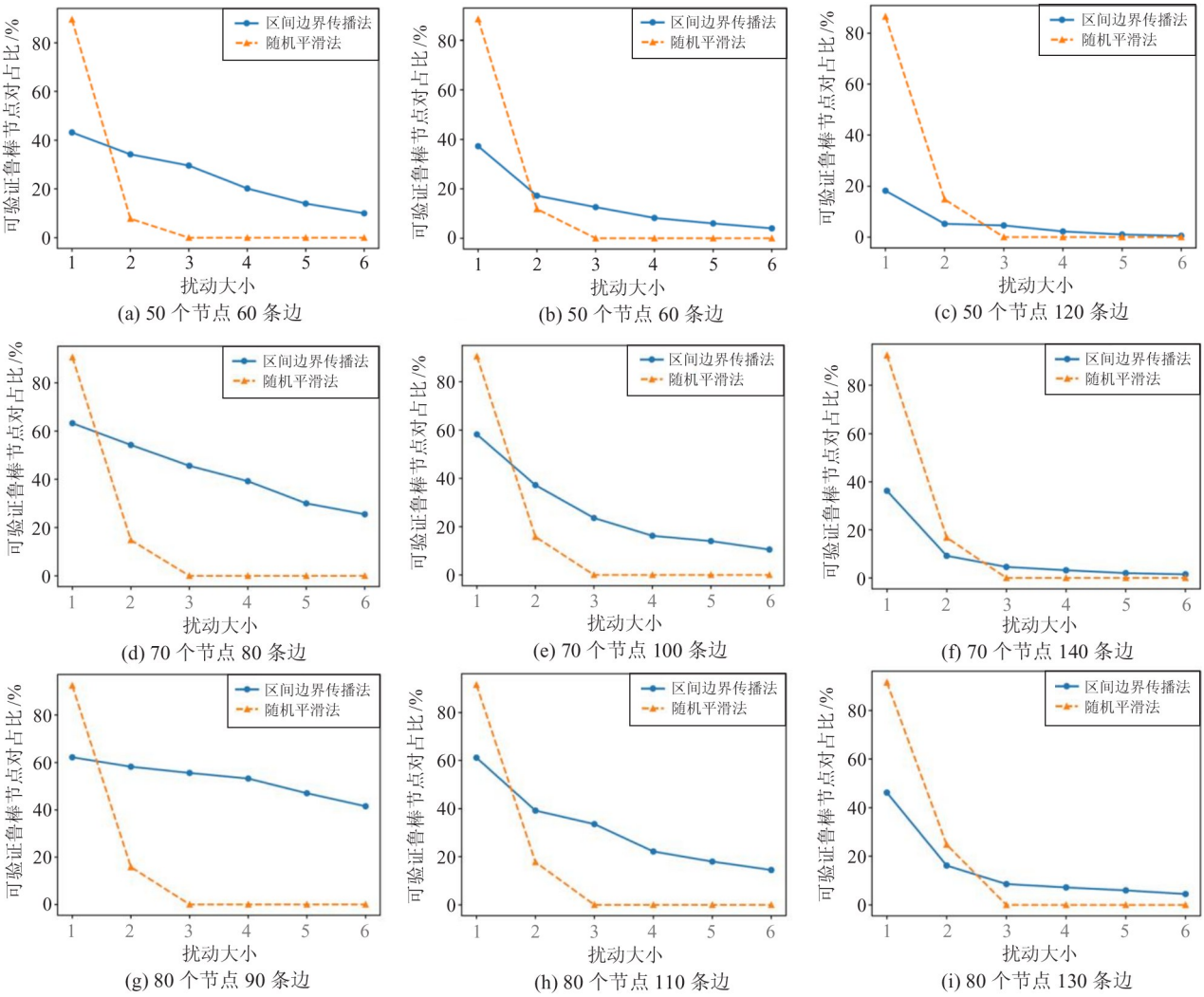

图 2　实验结果对比

3.3　参数分析

表 2 为在节点嵌入向量维数为 3 情况下的实验结果。可以观察到,当节点数固定时,随着边数的增加,可验证鲁棒节点对的占比呈下降趋势。例如,在节点数为 50 的情况下,边数从 60 增加到 150,可验证鲁棒节点对的占比从 52.52% 下降到 9.22%。同时,当边数确定,随着节点数的增加,可验证鲁棒节点对的占比呈上升趋势。例如,在边数为 90 的情况下,节点数从 50 增加到 80,可验证鲁棒节点对的占比从 33.41% 上升到 62.25%。这种趋势的变化可能与图网络的平均度有关。

表 2　不同节点数和边数下可验证鲁棒节点对占比

节点数 \ 边数	60	70	80	90	100	110	120	130	140	150
50	52.5%	43.19%	37.65%	33.41%	29.25%	20.46%	18.43%	12.68%	9.44%	9.22%
60	-	56.12%	53.41%	47.64%	42.16%	36.34%	30.83%	30.97%	23.73%	20.86%
70	-	-	63.63%	59.85%	58.23%	52.71%	50.89%	47.68%	36.56%	33.39%
80	-	-	-	62.25%	63.26%	61.72%	57.41%	51.67%	52.28%	46.96%

图 3 显示了在不同节点数下平均度与鲁棒验证率的关系。可以明显观察到,随着平均度的增大,鲁棒验证率呈下降趋势。这可能是因为随着图网络的平均度的增加,节点之间连接的增加导致网络结构更为复杂,能被扰动的范围更大,使得这种鲁棒性验证方式变得更加困难,最终导致本文提出的鲁棒性验证方法的效率下降。

表 3 关注边数为 100 时的实验结果,研究了不同嵌入向量维数对可验证鲁棒节点对占比的影响。

同时,如图 4 所示,随着嵌入维数的增加,可验证鲁棒节点对的占比呈下降趋势。例如,在节点数为 50 的情况下,当嵌入维数从 3 增加到 8 时,可验证鲁棒节点对的占比从 29.25% 下降到 4.61%。

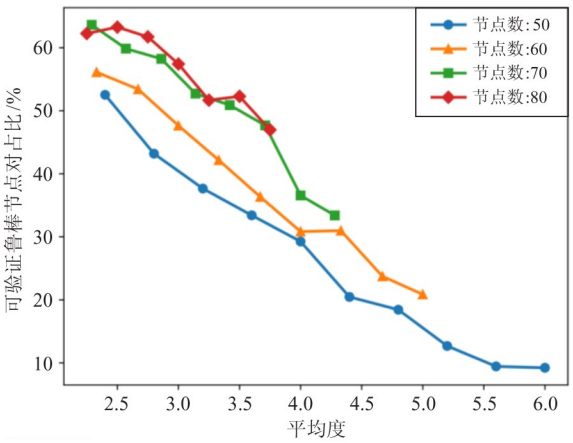

图 3　平均度对可验证鲁棒节点对占比的影响

表 3　不同嵌入维数和节点数下可验证鲁棒节点对占比

节点数 \ 维数	3	4	5	6	7	8
50	29.25%	27.04%	15.10%	11.72%	7.42%	4.61%
60	42.16%	41.03%	32.53%	22.80%	19.90%	14.95%
70	58.23%	51.68%	46.80%	37.85%	32.47%	25.28%
80	63.26%	61.65%	56.83%	49.85%	39.03%	34.20%

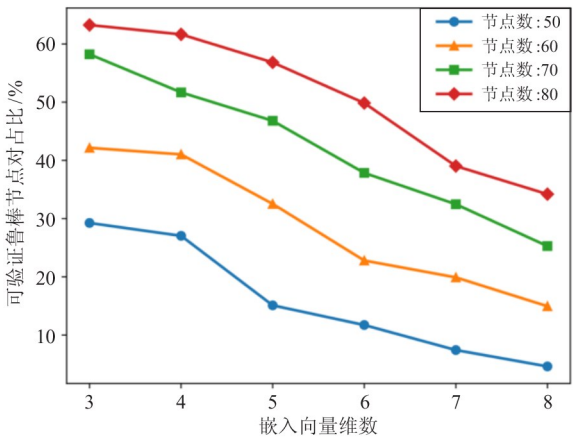

图 4　嵌入向量维数对可验证鲁棒节点对占比的影响

随着节点嵌入向量维数的增加,网络中每个节点的嵌入向量变得更为丰富和复杂。高维度的嵌入向量能够更准确地捕捉节点之间的结构和关系,提供更详细的网络表示。然而,这种提高也导致边界传播时的计算复杂性增加。在 IBP 中,边界的传播涉及到对输入空间中的不确定性进行推导,而高维度导致较大的输入空间,这使得边界更宽松,最终导致鲁棒验证率的下降。因此,在选择嵌入维数时,需要在鲁棒性验证效率和计算效率之间找到平衡点。

4　结　论

针对 DeepWalk 模型鲁棒性验证研究,本文提

出了一种鲁棒性验证方法,该方法以区间边界传播思想为基础,旨在评估在度值攻击下,基于 DeepWalk 算法的链路预测任务的稳定性与可靠性。通过实验分析观察到,在特定的网络参数条件下,此方法能够识别并计算出具有绝对鲁棒性的节点对。然而,随着网络平均度的提升以及节点嵌入向量维数的增加,可被验证为鲁棒的节点对所占比例呈现下降趋势。这一现象的原因可能在于网络结构的复杂性和嵌入空间的维度增长导致扰动空间增大,致使算法在处理鲁棒性验证时的难度增加。这一发现对于进一步优化鲁棒性验证算法具有理论价值。

综上所述,尽管本文实现了对 DeepWalk 的鲁棒性验证,但在更复杂的网络结构和高嵌入维度条件下如何优化鲁棒性验证算法需要更进一步的研究。同时,将这种验证方法泛化到其他图嵌入算法仍然是一个值得深入研究的方向。

参考文献

[1] Perozzi B, Al-Rfou R, Skiena S. Deepwalk: online learning of social representations[C]//Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, USA: ACM, 2014:701-710.

[2] Qiu J, Dong Y, Ma H, Li J, et al. Network embedding as matrix factorization: unifying DeepWalk, line, PTE, and node2vec[C]//Proceedings of the 11th ACM International Conference on Web Search and data Mining. Los Angeles, USA: ACM, 2018:459-467.

[3] 冶忠林,曹蓉,赵海兴,等. 基于矩阵分解的 DeepWalk 链路预测算法[J]. 计算机应用研究, 2020,37(2):424-429.

[4] Guimerà R, Sales-Pardo M. Missing and spurious interactions and the reconstruction of complex networks[J]. Proceedings of the National Academy of Sciences, 2009, 106(52):22073-22078.

[5] Qi G J, Aggarwal C, Huang T. Link prediction across networks by biased cross-network sampling[C]//IEEE 29th International Conference on Data Engineering. Brisbane, Australia: IEEE, 2013:793-804.

[6] Huang Z, Li X, Chen H. Link prediction approach to collaborative filtering[C]//Proceedings of the 5th ACM/IEEE-CS Joint Conference on Digital Libraries. Denver, USA: ACM, 2005:141-142.

[7] Song H H, Cho T W, Dave V, et al. Scalable proximity estimation and link prediction in online social networks[C]//Proceedings of the 9th ACM SIGCOMM Conference on Internet Measurement. Chicago, USA: ACM, 2009:322-335.

[8] Yu S Q, Zheng J, Wang J H, et al. Unsupervised Euclidean distance attack on network embedding[C]//2020 IEEE 5th International Conference on Data Science in Cyberspace. Hong Kong, China: IEEE, 2020:71-77.

[9] Dai Q, Li Q, Tang J, et al. Adversarial network embedding[C]//Proceedings of the AAAI Conference on Artificial Intelligence. New Orleans, USA: AAAI Press, 2018:2167-2174.

[10] Dai Q, Shen X, Zhang L, et al. Adversarial training methods for network embedding[C]//The World Wide Web Conference 2019. San Francisco, USA: International World Wide Web Conferences Steering Committee, 2019:329-339.

[11] Yang C, Shi Q. An interval perturbation method for singular value decomposition (SVD) with unknown-but-bounded (UBB) parameters[J]. Journal of Computational and Applied Mathematics. 2024,436:115436.

[12] Sunaga T. Theory of an interval algebra and its application to numerical analysis[J]. Japan Journal of Industrial and Applied Mathematics, 2009,26:125-143.

[13] Gowal S, Dvijotham K D, Stanforth R, et al. Scalable verified training for provably robust image classification[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Seoul, Korea: IEEE, 2019:4842-4851.

[14] Zügner D, Günnemann S. Certifiable robustness and robust training for graph convolutional networks[C]//Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Anchorage, USA: ACM, 2019:246-256.

[15] Zügner D, Günnemann S. Certifiable robustness of graph convolutional networks under structure perturbations[C]//Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Online: ACM, 2020:1656-1665.

[16] Wong E, Kolter Z. Provable defenses against adversarial examples via the convex outer adversarial polytope[C]//2018 International Conference on Machine Learning.

Stockholm, Sweden: ICML, 2018:5286-5295.

[17] Lecuyer M,Atlidakis V,Geambasu R,et al. Certified robustness to adversarial examples with differential privacy[C]// 2019 IEEE Symposium on Security and Privacy. San Francisco, USA: IEEE, 2019:656-672.

[18] Jia J, Wang B, Cao X, et al. Certified robustness of community detection against adversarial structural perturbation via randomized smoothing[C]//Proceedings of The Web Conference 2020. Taibei, China: ACM, 2020: 2718-2724.

[19] Wang B, Jia J, Cao X, et al. Certified robustness of graph neural networks against adversarial structural perturbation[C]// Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. Online: ACM, 2021:1645-1653.

[20] Levy O, Goldberg Y. Neural word embedding as implicit matrix factorization[C]//Advances in Neural Information Processing Systems 2014. Montreal, Canada: NeurIPS, 2014:27.

[21] Yang C, Liu Z. Comprehend DeepWalk as matrix factorization[EB/OL]. (2015-01-02)[2024-02-26]. https://arxiv.org/pdf/1501.00358.

[22] Yang C, Liu Z, Zhao D, et al. Network representation learning with rich text information[C]// Proceedings of the 24th International Joint Conference on Artificial Intelligence. Buenos Aires, Argentina: IJCAI, 2015:2111-2117.

Certified robustness of DeepWalk based on interval bound propagation

Yu Shanqing, Fang Xu, Wang Jinhuan, Peng Songtao

(College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023)

Abstract

The highly nonlinear and complex nature of graph embedding algorithms based on DeepWalk makes it challenging to certify their robustness. Therefore, this paper proposes a method based on interval boundary propagation (IBP) to certify the robustness of DeepWalk against degree value attacks in link prediction tasks. This paper equates DeepWalk to the task of singular value decomposition of the target matrix. Following the idea of IBP, it first propagates the perturbation range from the original degree value matrix to the target matrix, then to the node's embedding vectors, obtaining their perturbation intervals as constraints. Finally, it solves an optimization problem to obtain the robust certification results. Furthermore, the impacts of different graph scales, graph densities, and embedding dimensions on the proportion of certifiably robust node pairs are experimentally investigated. The experimental results demonstrate that the average robust certification rate can reach up to 63.63%.

Key words: DeepWalk, link prediction, interval bound propagation, certified robustness