

# 对抗性网络嵌入

**摘要：**网络的低维表示在诸如节点分类、链路预测和网络可视化等各种任务中被证明是有效的。现现有的方法可以有效地将不同的结构属性编码到表示中，如邻域连接模式、全局结构角色相似性和其他高阶相似性。然而，除了捕获网络结构属性的目标之外，它们中的大多数都缺乏增强表示的健壮性的额外约束。本文旨在利用生成型对抗性网络在捕捉潜在特征方面的优势，研究其在学习稳定鲁棒图表示方面的贡献。具体来说，我们提出了一个对抗性网络嵌入(ANE)框架，它利用对抗性学习原理来规范表示学习。它由两个部分组成，即结构保持部分和对抗性学习部分。前者的目的是捕捉网络结构属性，而后者则是为了捕获网络结构属性。通过将潜在表示的后验分布与给定的先验匹配来学习鲁棒表示。如实证结果所示，我们的方法是与或支持竞争的。基准网络嵌入任务的最新方法。源代码将在网上提供。

## 介绍

图是组织关系复杂的数据对象的一种自然方式，它对图中节点的丰富信息进行编码。例如，论文引文网络能够捕捉到创新流程的信息，并能再现论文之间的主题相关性。为了分析图，一种有效的方法是学习图中节点的低维表示，即节点嵌入(Perozzi, Al-Rfou和Skiena, 2014年; Tang等人。2015年; 曹、陆、徐2015)。学习到的表示应该编码有意义的语义、关系和结构信息，以便它们可以用作下游任务的功能，如网络可视化、链接预。词和节点分类。由于图形数据的高维性、稀疏性和非线性，网络嵌入是一个具有挑战性的研究课题。

近年来，人们提出了许多网络嵌入方法，如DeepWalk(Perozzi, AlRfou, Skiena 2014), line(Tang等)。2015年)和node2vec(Grover和Leskovec, 2016年)。他们的目标 在表示学习过程中捕获网络中的各种连接模式。这些模式包括局部邻域连通性、一阶和二阶近邻关系、全局关系。结构角色相似(即结构等价)和其他高阶近似值。如文献所示，在许多网络a中，网络嵌入方法更有效。分析的任务比一些经典的方法，如公共邻居(LibenNowell和Kleinberg, 2007年)和光谱聚类(Tang和Liu, 2011年)。

尽管现有方法在具有不同精心设计的目标的结构中是有效的，但是它们缺乏用于增强所学习的表示的鲁棒性的额外约束。当处理在实际应用中非常常见的嘈杂网络数据时，这些无监督的网络嵌入技术很容易导致表示不良。因此，在表征学习的过程中考虑一些不确定性是至关重要的。一种用于无监督方式的稳健表示学习的著名技术是去噪自动编码器(Vincent等人, 2010)。它通过从损坏的输入中恢复干净的输入来获得稳定且稳健的表示，即去噪标准。在(Cao, Lu和Xu 2016)中，作者将这一标准应用于网络嵌入。最近，许多生成对抗模型(Radford, Metz和Chintala 2016; Makhzani等。2016年 Donahue, Krahenbühl和Darrell 2017; 杜穆林等人。2017) 也被提议用于学习强大且可重复使用的表示。它们被证明在学习图像表示(Radford, Metz和Chintala 2016)和文本数据(Glover 2016)方面是有效的。但是，这些模型都没有专门设计用于处理图形数据。

在本文中，我们提出了一种称为Adversarial Network Embedding (ANE)的新方法，通过利用广告学习原理来学习鲁棒的网络表示(Goodfellow et al.2014)。除了优化保留结构的目标之外，还引入了对抗性学习过程来对数据不确定性进行建模。图1展示了著名的Zachary的空手道网络对敌对学习的影响。通过比较没有/具有对抗性学习正则化的两种方案的表示，可以容易地发现后一种方案获得更有意义和更健壮的表示。

更具体地说，ANE自然地将结构保留组件和对抗性学习组件组合在一个统一的框架中。前一个组件可以帮助捕获网络结构属性，而后者通过对抗来自某些先前分布的样本的对抗训练，有助于学习更强大的表示。对于结构保留，我们提出了适合我们的ANE框架的DeepWalk的归纳变体。它保持随机游走以探索节点的邻域并优化类似的目标函数，但采用参数化函数生成嵌入向量。此外，对抗性学习组件包括两部分，即生成器和鉴别器。它作为学习稳定和鲁棒特征提取器的正则化器，通过对抗训练对嵌入向量进行预先分配来实现。据我们所知，这是第一个用对抗性学习原理设计网络嵌入模型的工作。我们通过基准数据集的网络可视化和节点分类来实证评估所提出的ANE方法。定性和定量结果证明了我们方法的有效性。

## 相关工作

---

### 网络嵌入方法

近年来，已经提出了许多无监督的网络嵌入方法，根据它们使用的技术可以将其分为三组，即概率方法，基于矩阵因子分解的方法和基于自动编码器的方法。概率方法包括DeepWalk (Perozzi, Al-Rfou和Skiena 2014)，LINE (Tang等人2015)，node2vec (Grover和Leskovec 2016) 等。DeepWalk首先通过随机游走从原始图中获取节点序列，然后使用Skip-gram模型 (Mikolov等人2013) 通过将节点序列视为单词句来学习潜在表示。LINE尝试在两个单独的目标函数中保留一阶和二阶邻近，然后直接连接表示。在 (Grover和Leskovec 2016) 中，作者提出使用偏向随机游走来确定相邻结构，这可以在同质性和结构等价之间取得平衡。它实际上是DeepWalk的变体。

基于矩阵分解的方法首先预处理邻接矩阵以捕获不同类型的高阶邻接，然后分解处理后的矩阵以获得图嵌入。例如，GraRep (Cao, Lu和Xu 2015) 采用正向点相互信息 (PPMI) 矩阵作为预处理，基于对DeepWalk中k步随机游走和k步之间等价性的证明。概率转移矩阵。HOPE (Ou等人, 2016) 使用高阶邻近度量预处理具有高阶邻近度量的邻接矩阵的邻接矩阵，例如Katz Index (Katz 1953)，其可以帮助捕获非对称传递性质。M-NMF (Wang等人, 2017) 通过基于模块化的社区检测模型 (Newman, 2006) 学习可以很好地捕捉社区结构的嵌入。

Autoencoder是一种广泛使用的模型，用于学习高维数据的复杂表示，旨在保留潜在空间中尽可能多的信息，以便重建原始数据 (Hinton和Salakhutdinov 2006)。DNGR (Cao, Lu和Xu 2016) 首先计算PPMI矩阵，然后通过堆叠去噪自动编码器来学习表示。SDNE (Wang, Cui和Zhu 2016) 是堆叠自动编码器的变体，其在损失函数中添加约束以迫使所连接的节点具有类似的嵌入向量。在 (Kipf和Welling 2016) 中，作者提出了一种变分图自动编码器 (VGAE)，它使用图形卷积网络 (Kipf和Welling 2017) 编码器来捕获网络结构特性。与变分自动编码器 (VAE) (Kingma和Welling 2014) 相比，我们的ANE方法明确地规定了潜在空间的后验分布，而VAE仅假设先验分布。

### 生成对抗性网络

生成性对抗网络 (GAN) (Goodfellow等人, 2014) 是深度生成模型，其中框架由两部分组成，即生成器和鉴别器。GAN可以被表述为Minimal maximal confrontation game，其中生成器旨在将数据样本从一些先前的分布映射到数据空间，而鉴别器试图从真实数据中分辨出假样本。由于缺乏明确的推理结构，该框架不直接适用于无监督的表示学习。

正如现有工作所证明的，这个问题有三种可能的解决方案。首先，一些工作人员将一些结构整合到框架中进行推理，即将数据空间中的样本投射回潜在特征的空间，例如BiGAN (Donahue, Krahenbühl和Darrell 2017)，ALI (Dumoulin等人, 2017) 和EBGAN (Zhao, Mathieu和LeCun 2016)。这些方法可以在许多应用中学习强大的表示，例如图像分类 (Donahue, Krahenbühl, 和Darrell 2017) 和文档检索 (Glover 2016)。第二种方法是从鉴别器的隐藏层生成表示，如DCGAN (Radford, Metz和Chintala 2016)。通过采用分数跨度卷积层，DCGAN可以从用于监督任务的生成器和判别器网络中学习表达图像表示。第三个想法是使用对抗性学习过程来规范表达。一个成功的做法是Adversarial Autoencoder (Makhzani等人, 2016)，它可以从未标记的数据中学习强大的表示，而无需任何监督。

### 对抗性网络嵌入

---

在本节中，我们将首先介绍问题的定义和将要使用的符号。在此基础上，我们将对所提出的对抗性网络嵌入框架进行概述，并对每个组件进行详细的描述。

#### 问题定义与标注

网络嵌入是为了学习信息网络中节点的有意义表示。信息网络可以表示为 $G=(V, E, A)$ ，其中 $V$ 是节点集， $E$ 是边集 每一个表示一对节点之间的关系，而 $A$ 是一个加权邻接矩阵，它的条目量化了对应关系的强度。特别地， $A$ 中的每个参数的值在未加权图中是0或1，指定在两个节点之间是否存在边。给定一个信息网络 $G$ ，网络嵌入正在进行从节点 $v_i \in V$ 到低维向量 $\mathbf{u}_i \in R^d$ 的映射，形式格式如下： $f: V \mapsto U$ ，其中 $\mathbf{u}_i^T$ 是 $U$  ( $U \in R^{N \times d}, N = |V|$ )的第 $i$ 行， $d$ 是表示的维数。我们称 $U$ 表示矩阵。这些表示应编码网络的结构信息。

## 框架概述

在这项工作中，我们利用对抗性学习原则来帮助学习稳定和强大的表示。图2显示了所提出的对抗网络嵌入(ANE)框架，它主要由两个部分组成，即结构保留组件和对抗性学习组件。具体而言，保留组件的结构专用于将网络结构信息编码到表示中。这些信息包括本地邻域连接模式，全局结构角色相似性以及其它高阶邻近性。实现此组件有许多可能的替代方案。实际上，现有方法 (Perozzi, Al-Rfou和Skiena 2014; Tang等2015; Cao, Lu和Xu 2015) 可以被认为是结构保留模型，但是没有任何约束来帮助增强其稳健性表示。在本文中，我们提出了一种用于结构保持的归纳DeepWalk。它保持随机游走以探索节点的邻域并优化类似的目标函数，但采用参数化函数 $G(\cdot)$ 来生成嵌入向量。在训练过程中，直接更新 $G(\cdot)$ 的参数，而不是嵌入向量。此外，对抗性学习组件包括两部分，即生成器 $G(\cdot)$ 和鉴别器 $D(\cdot)$ 。它作为学习稳定和滚动特征提取器的正则化器，通过对抗训练对嵌入向量进行预先分配来实现。需要强调的是，参数化函数 $G(\cdot)$ 由结构保持组件和对抗性学习组件共享。这两个组件将在训练过程中交替更新 $G(\cdot)$ 的参数。

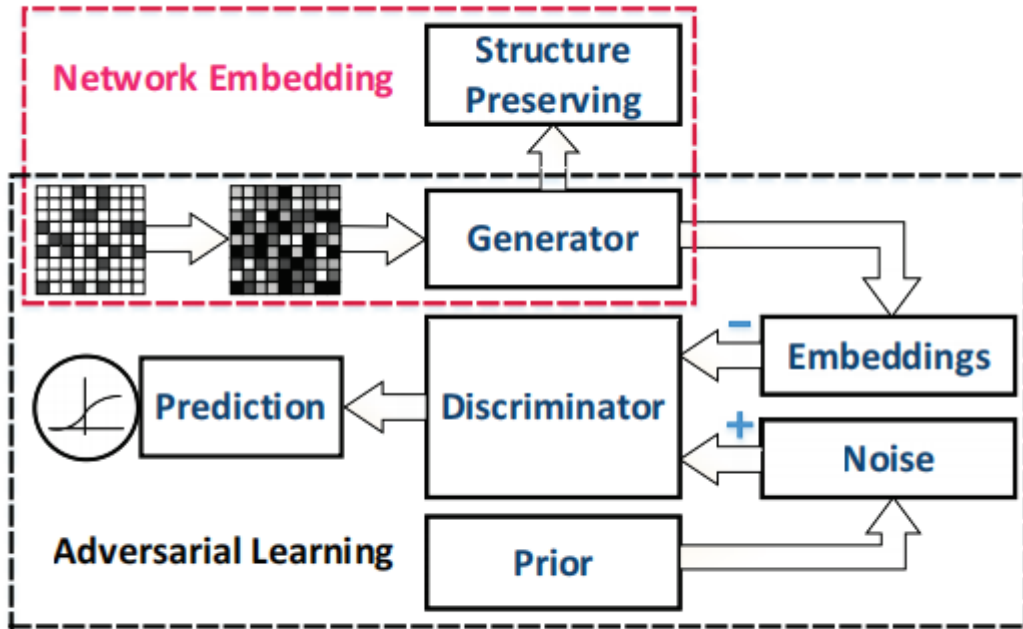


Figure 2: Adversarial Network Embedding Framework

## 图预处理

在实际应用中，信息网络通常非常稀疏，在训练深度模型时可能导致严重的过度拟合问题。为了帮助减轻稀疏性问题，一种常用的方法是预先建立具有高阶邻近性的邻接矩阵 (Tang et al.2015; Cao, Lu和Xu 2015) 。在本文中，我们将移位的PPMI矩阵 $X$  (Levy和Goldberg 2014) 作为生成器1的输入特征，定义为：

$$X_{ij} = \max \left\{ \log \left( \frac{M_{ij}}{\sum_k M_{kj}} \right) - \log(\beta), 0 \right\}$$

其中  $M = \hat{A} + \hat{A}^2 + \dots + \hat{A}^t$  可以捕获不同的高阶序列， $\hat{A}$  是在逐行归一化后从加权邻接矩阵  $A$  获得的1步概率转移矩阵，并且  $\beta$  设置为 本文中为  $\frac{1}{N}$ 。  $X$  中的行向量  $\mathbf{x}_i^T$  是表征图  $G$  中的节点  $v_i$  的上下文信息的特征向量，但具有高维度。

## 结构保持模型

理想情况下，现有的无监督网络嵌入方法可以用作我们框架中的结构保留组件，用于将节点依赖性编码为表示。然而，其中许多是转换方法，其嵌入查找作为嵌入式生成器，例如DeepWalk和LINE，由于我们使用参数化生成器作为标准GAN，因此它们不直接适用于对抗性学习组件的生成器。利用参数化的生成器，我们的框架可以很好地处理具有节点属性的网络，并利用深度学习模型探索网络的非线性特性。在这项工作中，我们设计了DeepWalk的归纳变体，适用于加权和未加权图形。从理论上讲，它也可以推广到具有节点属性的网络中看不见的节点，如某些归纳方法（Yang, Cohen和Salakhutdinov 2016; Hamilton, Ying和Leskovec 2017），但我们在本文中没有探讨它。此外，我们还研究使用去噪自动编码器（Vincent et al.2010）作为结构保留组件。

**Inductive DeepWalk (IDW)** IDW模型使用random walk来测量DeepWalk中的节点序列。从每个节点  $v_i$  开始， $\eta$  序列被随机抽取，长度为  $l$ 。在每个步骤中，从当前节点的邻居中选择一个新节点，其概率与矩阵  $A$  中的相应权重成比例。为了提高效率，采用别名表方法（Li et al.2014）进行采样 来自每个采样步骤中设置的候选节点的节点。在单个采样步骤中仅花费  $O(1)$  时间。然后，可以从节点序列构造正节点对。对于每个节点序列  $W$ ，我们将正目标 - 上下文对确定为集合  $\{(w_i, w_j) : |i - j| < s\}$ ，其中  $w_i$  是序列  $W$  中的第  $i$  个节点， $s$  表示上下文大小。

类似于Skip-gram（Mikolov等人，2013），节点  $v_i$  具有两个不同的表示，即目标表示  $\mathbf{u}_i$  和上下文表示  $\mathbf{u}'_i$ ，它们分别由目标发生器  $G(\cdot; \theta_1)$  和上下文生成器  $F(\cdot; \theta'_1)$  生成。生成器是参数化函数，在这项工作中使用神经网络实现。给定对应于节点  $v_i$  的  $X$  中的行向量  $\mathbf{x}_i^T$ ，我们得到  $\mathbf{u}_i = G(\mathbf{x}_i; \theta_1)$  和  $\mathbf{u}'_i = F(\mathbf{x}_i; \theta'_1)$ 。为了捕获网络结构属性，我们使用负抽样方法为每个正目标 - 上下文对  $(v_i, v_j)$  定义了跟随目标函数：

$$\mathcal{O}_{IDW}(\theta_1; \theta'_1) = \log \sigma \left( F(\mathbf{x}_j; \theta'_1)^T G(\mathbf{x}_i; \theta_1) \right) + \sum_{n=1}^K \mathbb{E}_{v_n \sim P_n(v)} \left[ \log \sigma \left( -F(\mathbf{x}_n; \theta'_1)^T G(\mathbf{x}_i; \theta_1) \right) \right]$$

其中  $\sigma(x) = 1/(1 + \exp(-x))$  是sigmoid函数， $K$  是每个正对的负样本数， $P_n(v)$  是采样负上下文节点的噪声分布， $\theta_1$  和  $\theta'_1$  是要学习的参数。如（Mikolov等人2013）所述， $P_n(v) = d_v^{3/4} / \sum_{v_i \in V} d_{v_i}^{3/4}$  在实践中可以达到相当好的性能，其中  $d_v$  是节点  $v$  的程度。

## 对抗性学习

使用对抗性学习组件来定期化表示。它由发生器  $G(\cdot; \theta_1)$  和鉴别器  $D(\cdot; \theta_2)$  组成。具体地， $G(\cdot; \theta_1)$  表示输入高维特征的非线性变换到嵌入矢量。 $D(\cdot; \theta_2)$  表示来自实际数据的样本的概率。生成器功能与保留组件的结构共享。与GAN不同（Goodfellow等人，2014），在我们的框架中，选择先验分布  $p(z)$  作为用于生成实际数据的数据分布，而嵌入向量被视为伪样本。在训练过程中，训练鉴别器以将先前样本与嵌入向量分开，而生成器旨在将嵌入向量拟合到先前分布。这个过程可以被认为是一个双人迷你极小游戏，发生器和鉴别器相互对战。鉴别器的效用函数是：

$$\mathcal{O}_D(\theta_2) = \mathbb{E}_{z \sim p(z)} [\log D(z; \theta_2)] + \mathbb{E}_{\mathbf{x}} [\log(1 - D(G(\mathbf{x}; \theta_1); \theta_2))]$$

为了将其输出伪装成先前的样本，对发生器进行训练以改善以下收益：

$$\mathcal{O}_G(\theta_1) = \mathbb{E}_{\mathbf{x}} [\log(D(G(\mathbf{x}; \theta_1); \theta_2))]$$

我们认为，对抗性学习组件可以帮助改善学习表征的稳健性和结构意义。我们用两个结构保留模型实例化我们的框架，即归纳深度漫游（IDW）和去噪自动编码器（DAE）。我们将带有IDW的ANE框架称为Adversarial Inductive Deep-Walk（AIDW），以便于说明。实际上，以DAE作为结构保留组件，ANE框架将成为一种对抗性自动编码器（Makhzani et al. 2016），我们将其表示为ADAЕ，以突出去除学习表示中的去噪标准的重要性。

在对抗性学习期间，选择适当的先前分布也很重要。像GANs研究中的许多实践（Radford, Metz和Chintala 2016; Makhzani等人2016; Donahue, Krahenbühl和Darrell 2017），先前的分布通常被定义为均匀或高斯噪声，使GAN能够学习有意义的和针对不确定性的强有力的表述。在我们的实验中，我们还考虑了具有两种先验分布的ANE框架，但没有发现显著差异。一个可能的原因是两种不确定性都可以帮助ANE框架工作以达到一定程度的抗噪声鲁棒性。在先前的领域知识的指导下仔细选择先前的分布可能会进一步改善特定于应用程序的性能。

## 算法

为了实施ANE方法，我们考虑联合列车运行程序，包括两个阶段，包括结构预先服务阶段和对抗性学习阶段。在结构保留阶段，我们优化AIDW的目标函数（2）。在对抗性学习阶段，通过极小极大优化问题对表示施加先验分布。首先，训练鉴别器以区分先前样本和嵌入向量之间的差异。然后，更新生成器的参数以使嵌入层叠向量适合于先前空间以欺骗鉴别器。可以采用（Arjovsky, Chintala和Bottou 2017）中提出的一些技巧来帮助提高学习的稳定性并避免传统GAN培训中的模式崩溃问题。

Table 1: Statistics of datasets

Dataset	$ V $	$ E $	$ C $
Cora	2,708	5,278	7
Citeseer	3,264	4,551	6
Wiki	2,363	11,596	17
Cit-DBLP	5,318	28,085	3

## 实验

### 实验设置

**数据集：**我们在四个真实数据集上进行实验，统计数据如表1所示，其中C表示标签集。Cora和Citeseer是由（McCallum等人，2000）构建的纸引文网络。Wiki（Sen et al.2008）是一个网络，节点作为网页，边缘作为网页之间的超链接。我们将这三个网络视为无向网络，并通过删除自循环和零度节点对原始数据集进行一些预处理。Cit-DBLP是一个从DBLP数据集中提取的纸质引文网络（Tang et al.2008）。

**基线：**我们将我们的模型与几种基线方法进行比较，包括DeepWalk、Line、GraRep和node2vec。有许多其他的网络嵌入方法，但是我们在这里不考虑它们，因为它们的性能不如相应的文献所示的这些基线模型。基线的说明如下。

- DeepWalk（Perozzi, Al-Rfou和Skiena 2014）：DeepWalk首先通过截断随机游走将网络转换为节点序列，然后将其用作Skip-gram模型的输入以学习表示。
- LINE（Tang et al.2015）：LINE可以通过建模节点共现概率和节点条件概率保留无向图的第一阶和第二阶邻域。
- GraRep（Cao, Lu和Xu 2015）：GraRep通过构造不同的k步概率转移矩阵来保留节点邻近性。
- node2vec（Grover和Leskovec 2016）：node2vec开发了一个有偏差的随机游走程序来探索节点的邻居，它可以在本地属性和网络的全局属性之间取得平衡。

**参数设置：**对于LINE，我们遵循参数设置（Tang et al.2015）。嵌入向量通过L2范数归一化。此外，我们通过向低度节点添加两跳邻居来专门预处理原始稀疏网络。对于DeepWalk和node2vec，窗口大小 $s$ ，步行长度 $l$ 和每个节点的步数 $\eta$ 分别设置为10,80和10，以进行公平的比较。对于GraRep，最大矩阵转换步骤设置为4，其他参数的设置遵循（Cao, Lu和Xu 2015）中的设置。请注意，所有方法的表示维度都设置为128，以便进行公平比较。

对于我们的方法，我们只使用最简单的发生器结构。具体来说，发电机是一个单层网络，在输出上具有泄漏的ReLU激活（泄漏0.2）和批量归一化（Ioffe和Szegedy 2015）。移位的PPMI矩阵 $X$ 是通过将 $Co$ 设置为4和Citeseer获得的，以及将3设置为Wiki获得的。对于感应式DeepWalk，负样本数 $K$ 设置为5，其他参数设置与DeepWalk相同。对于去噪自动编码器，它只有一个隐藏层，其维数为128。对于框架的鉴别器，它是一个三层神经网络，层结构为512-512-1。对于前两层，我们使用泄漏的ReLU激活（泄漏为0.2）和批量标准化。对于输出层，我们使用sigmoid激活。对于AIDW和ADAe，结构保留组件的设置分别与IDW和DAE的设置相同。对抗性学习组件的先验分布设置为 $z_i \sim U[-1, 1]$ 。我们使用RMSProp优化器，学习率为0.001。

## 网络可视化

网络可视化是分析高维图形数据不可或缺的方法，可以直观地揭示数据的内在结构（Tang et al.2016）。在本节中，我们使用t-SNE可视化由几个不同模型生成的节点的表示（van der Maaten和Hinton 2008）。我们从DBLP构建了一个纸质引文网络，即Cit-DBLP，来自三个不同的出版部门，包括信息科学，ACM图形交易和人机交互。表1中列出了该数据集的一些统计数据。这些论文基于它们所属的研究领域自然分为三类。

图3显示了使用t-SNE在相同参数配置下从不同模型获得的嵌入向量的可视化。对于DeepWalk和LINE，来自不同类别的纸张在图的中心彼此混合。对于LINE，有6个聚类，每个类别对应两个独立的聚类，这与网络的真实结构相冲突。此外，不同集群之间的界限尚不清楚。node2vec和IDW的可视化形成了三个主要集群，这些集群优于DeepWalk和LINE。但是，node2vec的蓝色集群和绿色集群之间的界限尚不清楚，而红色集群和绿色集群的边界对于IDW来说有点混乱。与基线方法相比，AIDW表现更好。我们可以观察到AIDW的可视化具有三个彼此之间具有相当大的余量的聚类。此外，每个簇可以与图中的另一个簇线性分离，这是图中所示的其他基线所不能实现的。直觉上，这个实验表明，广告学习正规化可以帮助学习更多有意义和更强大的表征。

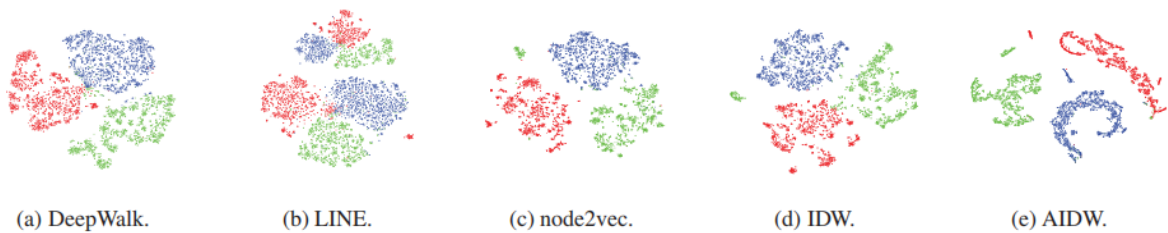


Figure 3: Visualization of Cit-DBLP dataset. Each point represents one paper. Different colors correspond to different publication divisions. Red: “Information Science”, blue: “ACM Transactions on Graphics”, green: “Human-Computer Interaction”.



Table 2: Accuracy (%) of multi-class classification on Cora

%Labeled Nodes	10%	20%	30%	40%	50%	60%	70%	80%	90%
DeepWalk	71.43	73.83	75.61	76.92	77.79	77.78	78.47	79.17	79.04
LINE	71.26	74.50	76.04	76.81	77.68	77.99	78.46	79.00	79.11
GraRep	74.78	76.78	78.56	78.99	79.39	79.85	79.96	80.94	81.29
node2vec	75.06	78.49	80.06	80.94	81.52	82.07	82.39	83.28	83.17
DAE	75.21	78.07	79.39	80.51	80.91	81.41	82.36	83.23	83.32
ADAE	75.01	77.45	79.65	80.96	81.64	82.11	82.62	83.52	84.10
IDW	66.32	72.21	75.23	76.65	77.66	78.32	79.20	79.93	80.63
AIDW	<b>76.93</b>	<b>79.50</b>	<b>81.31</b>	<b>82.01</b>	<b>82.28</b>	<b>83.03</b>	<b>83.23</b>	<b>84.46</b>	<b>84.21</b>

Table 3: Accuracy (%) of multi-class classification on Citeseer

%Labeled Nodes	10%	20%	30%	40%	50%	60%	70%	80%	90%
DeepWalk	49.45	52.68	54.60	55.71	56.44	57.04	57.42	58.04	59.11
LINE	47.70	51.04	52.95	54.19	55.00	55.82	56.02	57.08	57.52
GraRep	51.62	53.29	53.59	53.83	54.55	54.62	54.97	54.90	56.27
node2vec	52.50	55.47	56.66	57.70	58.81	59.26	60.10	60.34	60.58
DAE	51.08	54.54	55.84	56.50	57.47	58.30	58.50	59.19	60.15
ADAE	52.40	55.58	56.64	57.32	58.34	59.60	60.27	60.63	61.31
IDW	45.45	50.47	52.32	53.38	54.75	55.18	55.98	56.23	57.19
AIDW	<b>53.25</b>	<b>56.76</b>	<b>57.95</b>	<b>59.06</b>	<b>59.45</b>	<b>59.95</b>	<b>60.28</b>	<b>60.87</b>	<b>62.26</b>

Table 4: Accuracy (%) of multi-class classification on Wiki

%Labeled Nodes	10%	20%	30%	40%	50%	60%	70%	80%	90%
DeepWalk	57.23	61.22	63.53	64.68	65.96	66.24	67.17	68.63	68.69
LINE	56.29	61.63	63.98	65.43	66.25	67.04	67.94	68.86	68.61
GraRep	<b>57.68</b>	61.14	62.73	63.89	64.86	65.55	66.01	67.55	68.02
node2vec	57.61	61.52	63.47	64.83	65.54	66.16	67.17	68.44	68.69
DAE	57.08	61.63	63.71	65.20	66.84	67.41	67.91	69.03	69.45
ADAE	57.24	61.67	63.85	65.34	66.67	67.11	67.79	69.68	70.59
IDW	56.01	60.77	63.08	64.37	65.66	66.47	67.15	67.86	68.52
AIDW	57.43	<b>62.14</b>	<b>64.18</b>	<b>65.53</b>	<b>67.07</b>	<b>68.00</b>	<b>69.44</b>	<b>71.63</b>	<b>72.03</b>

## 节点分类

标签信息可以指示节点的兴趣，信念或其他特征，这可以帮助促进许多应用，例如在线社交网络中的朋友推荐和目标广告。但是，在许多真实的世界环境中，只标记了一部分节点。因此，可以进行节点分类以挖掘未标记节点的信息。在本节中，我们对三个基准数据集进行了多类分类，即Cora，Cite? seer和Wiki。我们将培训比率从10%到90%进行综合评估。所有实验均使用Liblinear package2中的支持向量分类器进行（Fan等人，2008）。

### 结果与讨论

为了保证实验结果的可靠性，我们采用了表2、表3和表4所示的10次平均实验结果。我们有以下意见：

- IDW使用DeepWalk在Cora和Wiki上产生类似的结果，在Citeseer上的性能稍差。建议的模型AIDW建立在IDW之上，具有附加的对抗性学习组件。它在所有训练比率的三个数据集上始终优于IDW和DeepWalk。例如，在Cora上，AIDW在整个训练比率设置中比DeepWalk的准确度提高了4%以上。它表明，adversarial学习正则化可以显着提高学习代表的稳健性和辨别力。定量结果还验证了我们之前在网络可视化分析中的定性结果
- 当将培训比率从10%变为90%时，ADAE在Citeseer上的准确度提高了大约1%，对Cora的结果稍微好一点，在Wiki上的性能也相当。它表明，在构建DAE时，ANE框架还可以指导学习更强大的嵌入。但是，我们注意到ADAE没有像AIDW那样在对应的结构保留模型上取得明显的改进。一个原因是去噪标准已经有助于学习稳定和稳健的表示（Vincent et al.2010）。
- 总的来说，建议的方法AIDW始终如一地执行所有基线。如表2,3和4所示，node2vec平均比DeepWalk，LINE和GraRep产生更好的结果。我们的方法可以进一步实现对node2vec的改进。更具体地说，AIDW在不

同训练比率的所有三个基准数据集上实现了最佳分类准确度，在Wiki上只有一个例外，训练比率为10%。

## 模型灵敏度

在本节中，我们将研究AIDW w.r.t参数的性能以及Cora数据集上的先验类型。具体来说，对于参数灵敏度分析，我们检查表示维度 $d$ ，步长 $l$ 和上下文大小 $s$ 如何影响节点分类的性能，训练比率为50%。请注意，除了要测试的参数外，所有其他参数都设置为默认值。我们还比较了AIDW与两种不同先验的性能，即高斯分布( $\mathcal{N}(0, 1)$ )和均匀分布( $U[-1, 1]$ )。

图4 (a) 显示了尺寸 $d$ 测试的结果。当尺寸从8增加到512时，精度首先表现出明显的增加，然后一旦尺寸达到128左右就趋于饱和。此外，性能是什么？AIDW对步行长度和上下文大小不敏感。如图4 (b) 所示，精度略有提高，然后当步长从40变为100时变得稳定。随着上下文大小的增加，性能先保持稳定，然后稍微降低如果上下文大小超过6，则如图4 (c) 所示。由于Cora的平均节点程度仅为1.95左右，因此由大的上下文大小引入的嘈杂的邻域引起降级。

图5显示了Cora的多级分类结果，训练比率从10%到90%不等。具有均匀先验的AIDW的精度曲线几乎与具有高斯先验的AIDW的精度一致。它表明，两种类型的先验可以有助于学习稳健的代表，没有显著差异。

## 结论

---

提出了一种学习鲁棒图表示的对抗性网络嵌入框架。该框架由一个结构保持组件和一个对抗性学习组合组成。新代币。在结构保持方面，我们提出了归纳深度行走来捕获网络结构特性。对于对抗性学习，我们提出了一个极小极大优化问题来施加先验分布。增强了表示的鲁棒性。通过对网络可视化和节点分类的实证评价，验证了该方法的有效性。