**dyngraph2vec: Capturing network dynamics using dynamic graph representation learning**

Goyal, P., Chhetri, S. R., & Canedo, A. (2019). dyngraph2vec: Capturing network dynamics using dynamic graph representation learning. *Knowledge-Based Systems*.

**Abstract**

学习图的表示是在向量空间中获取图的各种性质的基本任务。最近的方法都是在学习静态网络中的表示。然而，真实世界的网络会随着时间的推移而发展，是具有动态性的。捕捉这种演化是预测未知网络特性的关键。为了了解网络动态对预测性能的影响，我们提出了一种嵌入方法，该方法学习了动态图的演化结构，可以对不可见的链路进行更高精度的预测。我们的模型dyngraph2vec使用由密集和重复层组成的深层体系结构来学习网络中的时间转换。我们的动机是需要捕捉动态，以预测使用随机块模型创建的玩具数据集。然后，我们在两个实际数据集上演示了dyngraph2vec相对于现有的先进方法的有效性。我们观察到，学习动态可以提高嵌入的质量，并产生更好的链接预测性能

**Introduction**

在这项工作中，我们的目标是捕获进化的潜在网络动态。给定图的时间快照，我们的目标是在每个时间步上学习节点的表示，同时捕获动态，以便预测它们未来的连接。

dyngraph2vec使用多个非线性层来学习每个网络中的结构模式，此外，它使用递归层来学习网络中的时间转换。重复层中的回看参数控制学习时间模式的长度

**贡献**：

1.提出了一种动态图嵌入模型dyngraph2vec，该模型能够捕获时间动态

2.我们证明了捕获网络动态可以显著提高链路预测的性能

3. 我们提出我们的模型的变化，以显示关键的优势和差异

**Related work**

图表示学习技术大致可分为两类: (i)静态图嵌入，用单个向量表示图中的每个节点，(ii)动态图嵌入，考虑图的多个快照，得到每个节点的向量时间序列。

**2.2 Dynamic graph embedding**

嵌入动态图是一个新兴的研究课题。提出了通过增加正则化来扩展静态图嵌入方法。

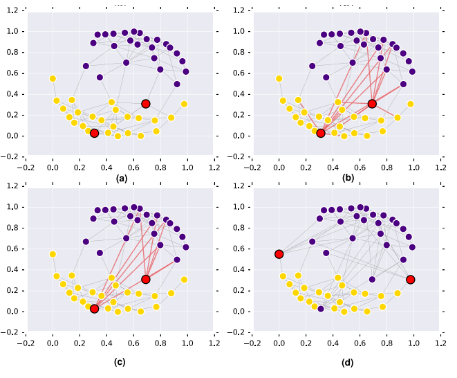
链接嵌入使得该方法对于高密度的图数据不可伸缩。我们的模型使用递归层来学习长序列图上的时间模式，并在每个时间步上使用多个全连接层来捕获复杂的模式。

**2.3 Dynamic link prediction**

在不强调图嵌入的情况下，提出了几种动态链路预测方法。这些方法中有许多使用了概率非参数方法

**3. motivation example**

考虑图G的演化，g={G1,…,GT}，Gt表示图在t时刻的状态。初始图G1由2个社区的随机块模型生成(图3中用靛蓝色和黄色表示)，块内和块间的概率分别设置为0.1和0.01。



演化模式可以定义为一个三步过程。

第一步：(如图3(a)所示)，我们从黄色代表的社区随机均匀地选取10个节点(图3中的红色部分显示了其中的两个节点)。

第二步：(如图3(b)所示)，我们在步骤1中选择的每个节点与Indigo社区中的随机节点之间随机添加30条边

第三步：(如图3(c)所示)中，步骤2中选择的节点的community membership由黄色更改为靛蓝色。根据社区内或者社区间的概率来删除边。

然后，进入下一个时间步骤(如图所示)图3(d))，重复相同的三个步骤来演化图形。非正式地说，这可以被解释为用户从一个社区到另一个社区的两步移动，首先增加另一个社区的朋友，然后移动到另一个社区。

我们的任务是学习对10个节点的社区变化的嵌入预测。图2显示了目前最先进的动态图嵌入技术(DynGEM、optimalSVD和DynamicTriad)的结果，以及我们的模式的三种变体。图2为第一步演化后节点的嵌入情况。为社区转移选择的节点用红色表示。从图2(a)可以看出，DynGEM使红色节点更靠近黄色群落的边缘，但是没有将任何节点移动到其他群落。类似地，图2(c)中的DynamicTriad结果表明，在下一步中，它只将1到4个节点转移到它的实际社区。图2(b)中optimal SVD方法不能移动任何节点

**4. methodology**

**4.1problem statement**

给定一个带权图G=(V,E)，我们用A表示G的邻接矩阵，Aij表示在图G中边(I,j)的权重，如果没有边ij，则Aij为0。G的演进可以表示为G={G1,…,GT}，Gt表示在t时刻图G的状态。

我们对问题的定义如下:给定一个演进图G，我们的目标是对每个节点v用一系列低维向量空间进行重新表示，其中yvt是节点v在时间t的嵌入。通过学习映射

和可以使得yvt捕获到时间模式并且用来预测yvt+1。换句话说，每个时间步长的嵌入函数使用来自图进化的信息来捕获网络动态，从而能够以更高的精度预测链接。

**4.2dyngraph2vec**

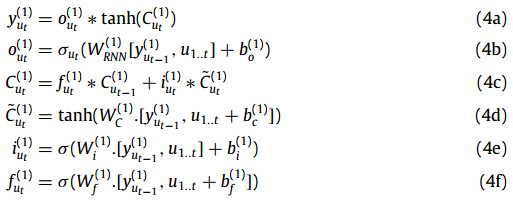
我们的dyngraph2vec是一个深度学习模型，它以一组之前的图形作为输入，并在下一个时间步生成图形作为输出，从而在每个时间步和跨多个时间步捕获顶点之间的高度非线性交互，由于嵌入值捕获了链接的时间演化，因此它允许我们预测下一步图链接。模型通过优化以下损失函数，在第t步学习网络嵌入:



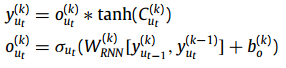
在这里，我们通过使用在t+l时刻的嵌入来惩罚在t+l+1时刻重建边缘的错误。最小化这个损失函数可以对参数进行调优，这样它就可以捕获节点之间的演化模式关系，从而在未来的时间步长中预测边缘。我们使用一个加权矩阵B来对观察到的比未观察到的链接更高的重建边进行加权，这与文献[19]中传统的方法相同。在这里对于边(I,j)属于Et+l+1。beta是一个超参数用来控制惩罚边的参数。

基于深度学习模型的架构，作者提出了三种模型变体，如图4所示: (i) dyngraph2vecAE, (ii) dyngraph2vecRNN,(iii) dyngraph2vecAERNN。我们的三种方法在函数f(.)的表达式上是不同的。

为了减少模型参数的数量，可以更有效的学习时间信息，作者提出了dyngraph2vecRNN和dyngraph2vecAERNN。在dyngraph2vecRNN中，我们使用稀疏连接的长短时记忆(LSTM)网络来学习嵌入。LSTM是一类递归神经网络(RNN)能够处理长期依赖问题。在动态图中，可能存在完全连接的自动编码器无法捕获的长期依赖关系。定义单个LSTM网络的隐藏状态表示为:



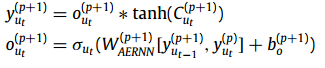
其中Cut表示LSTM中的cell状态，fut是触发forget-gate的值，out是触发输出门的值，iut表示触发LSTM更新门的值，C˜ut代表了新的估计候选状态，b代表偏差。可以在第一层连接l个LSTM网络。从t−l到t LSTM网络的链中，单元状态和隐藏表示在哪里传递。第k层表示如下:



当稀疏邻域向量被简化为低维表示时，LSTM网络可以更好地学习时间表示。为此，我们提出了dyngraph2vec模型的一个变体dyngraph2vecAERNN。在dyngraph2vecAERNN中，我们没有传递稀疏的邻域向量，而是使用一个全连接编码器来初始获取低维的隐藏表示，具体如下:



其中p为全连接编码器的输出层。然后将这个表示传递给LSTM网络。



然后将由LSTM网络生成的隐藏表示传递给一个完全连接的解码器

4.3 optimization

对上述损失函数进行了优化，得到了最优的模型参数。通过对式(1)中的解码器权值应用梯度，可以得到：

