**Influential Nodes Detection in Dynamic Social Networks**

Hafiene N, Karoui W, Romdhane L B. Influential Nodes Detection in Dynamic Social Networks[C]//International Conference on Business Information Systems. Springer, Cham, 2019: 62-73.

**1. introduction**

社交网络看起来像一个由节点和边组成的图结构，节点代表用户，边代表连接用户之间的交互，用户发展他们彼此之间的联系，而他们之间的关系随着时间而变化。一个不断变化的社交网络可以由各个时间点上的时间戳{G1,G2,…Gn}来观察，它不仅包含节点之间的一组关系，还包含关于这些关系如何在每个时间戳中更改的信息

我们提出了一个扩展的SND来解决在不断变化的社交网络下重要节点检测的问题。

**2.related work**

本节讨论对动态社交网络的不同研究的回顾。在本文中，我们主要研究了边缘变化的社交网络中影响节点的检测问题。

**2.1 Methods Based on a Non-linear Model**

Aggarwal等人提出了[7]，这是第一篇研究基于临时交互的动态网络中信息流权限确定问题的论文，此外，本文还研究了影响传播问题和从给定的传播模式中回溯影响节点的问题。他们考虑了如何在一段时间内发现一组影响最大的节点，他们将影响扩散模型建模为非线性系统，这与线性阈值模型或独立级联模型等触发模型有很大不同。[7]中的算法是启发式的，产生的结果没有任何可证明的质量保证。

**2.2 Methods Based on Metrics**

近年来，有人提出了一些中心度指标来估计节点的重要性，度中心性、中介中心性和亲密中心性是识别影响节点的三种基本中心性指标。然而，所有这些中心性度量都是为静态网络设计的，由于这个不足，为了将这些指标应用到不断变化的网络中，我们添加了关于边缘演化时间的信息。Kitsak等人[1]提出了k-shell分解，他们发现影响节点是位于网络核心的节点。k-shell分解在同一个k-shell中分配多个节点。

**2.3 Methods Based on a Diffusion Model**

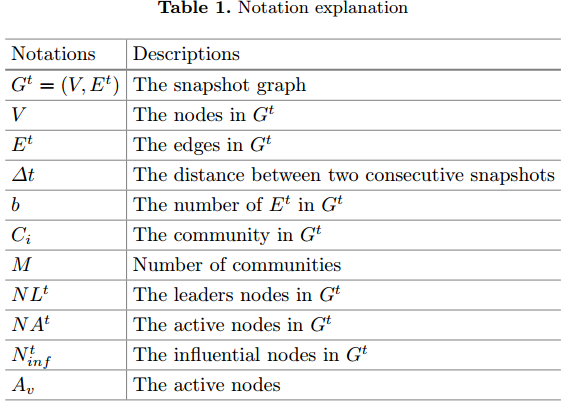
Chen等人[8]在影响扩散模型中加入了时间，用来跟踪在不断变化的社交网络中有影响力的节点，主要的思想是选择一个子集使其在ϕ个时间戳后可以最大化传播影响。然而，他们假设动态网络是完全观察到的。这在许多实际情况下是不可能的。Ohsaka等人[16]研究了一个相关的问题，在IC模型下对网络更新流维护一些RR集，使近似影响最大化能够以固定的概率完成。Tong等人[13]提出了动态独立级联(Dynamic Independent Cascade, DIC)模型，通过扩展经典的IC模型，DIC能够更好地控制真实社交网络的动态方面。

**3. Preliminaries and Problem Statement**

在本节中，我们首先介绍静态网络的扩散模型。然后，作者提出了一个扩展算法，作为动态社交网络中影响节点检测问题的推广

**3.1 problem statement**

之前的SND算法只针对静态社交网络检测有影响的节点。然而，真实的社交网络是动态的，所以无论是结构还是与边缘相关的影响力传播都在不断变化。因此，根据网络结构的演化和影响传播，需要改变影响传播最大化的领导节点。在本文中，我们将动态社交网络模型化为一组快照图，其中节点保持不变，而每个快照图中的边在不同的时间间隔内发生变化。我们将快照图表示为Gt=(V,Et)，其中V={v1,…vi,…vn}。我们的主要思想是识别一组领导节点，记作NLt= 1,…,T，这使得快照图Gt中的影响传播最大化。每个用户都由一个属性向量表示，其中xij是节点vi在属性j的数值。在SND中这个数值是二进制的，如果用户喜欢兴趣中心，值为1，否则为0。这种方法一方面利用了网络顶点之间的关系，另一方面又利用了它们的特性。



**4. proposed algorithm**

作者的算法包含两个阶段。事实上，SNDUpdate的第一个阶段探索了网络的结构方面，第二个阶段集中于语义方面。

**4.1 Phase 1: Community Detection**

在这个阶段，我们建议使用Combo[11]。该算法处理社区检测问题，能够处理各种目标函数。这个算法可以合并社区，将一个社区划分为两个，在两个不同的社区之间移动节点。

**4.2 Phase 2: Influential Nodes Detection**

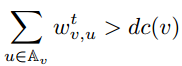
阶段2的第一部分支持生成一组leader节点。如果节点的度中心性大于或等于它的邻居，则节点是leader。度中心性允许度量节点与其邻居的连接总数。本文将动态社交网络模型化为一组快照图，每个快照图的边在不同的时间间隔内发生变化，从而导致每个快照图的leader节点发生变化，下面的方程用于计算给定快照图节点v的度中心性

。

在本阶段的第二部分中，我们使用[2]中的扩散模型来定义可以激活的非活动节点。在Gt中两个节点之间的链接与权重相关联，权重由它们的信息的语义相似性定义。两个节点之间在各快照图Gt中的语义相似度如下式所示:

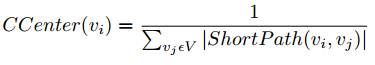


分区中的每个节点v与中心度有关，考虑到Gt中的每个节点和一组活动节点的中心性程度，如果其活动邻居的相似性总和或总权重超过了与节点v关联的阈值dc(v)，则节点v变为活动的。作者的扩散模型定义如下：



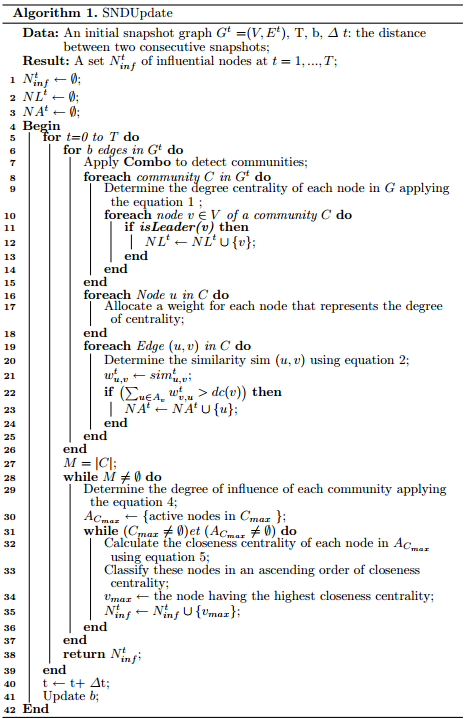
当我们在每个快照Gt中应用我们的扩散模型时，我们可以得到一组活动节点NAt，它允许我们识别一组有影响力的节点。我们的主要目标是从活动节点中识别出影响传播最大的节点。在每个社区中，我们确定其影响程度，由式4给出，等于每个社区中活跃节点的数量(Vactive)除以图G(V)中节点的总和。

  
首先，该方法确定了社区中影响程度最大的节点。其次，我们的目标是计算每个活动节点的贴近度中心性，由式5给出。它表示节点与图中所有其他节点之间最短路径的总长度



一旦我们计算出每个leader节点的close centrality，作者就会根据递增的顺序对这些节点进行分类。最具影响力的节点是接近中心性最高的节点。

SNDUpdate算法如下图所示：



首先，利用Combo算法对社区进行检测。

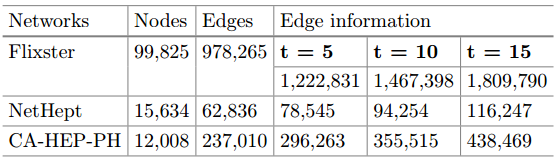
其次，在每个快照图Gt上，生成一组leading节点。

最后，应用扩散模型确定活动节点集，并确定影响力最大的节点

**5.实验**

**5.1 实验设定**

在我们的实验中，我们使用了三个真实的社交网络，如表2所示



对于每个数据集，我们使用了3个不同的社交网络，更新边缘，从而执行了3次实验。对于这3个示例，尽管原始网络和更新后的边缘不同，但是这3组快照的组合与数据集本身相似。用b表示在每个快照图中边的数量,使用不同的参数b和Δt,我们可以创建一组快照与许多属性的下一个实验。

**5.2 实验结果**

在本文中，作者他们的算法与两种静态影响最大化算法IMM，degree和一种动态算法DIM进行了比较

