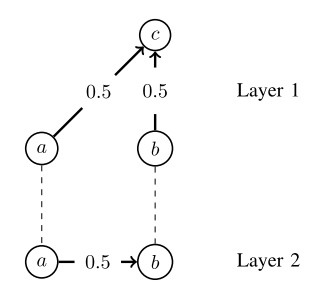
**Multiplex influence maximization in online social networks with heterogeneous diffusion models**

**1. Introduction**

越来越多的人使用在线社交网络，在在线社交网络中进行商品推广是个不错的选择。用户现在基本都在多个osn上有账户，因此，信息的传播可以通过这些重叠的用户从一个OSN跨到另一个OSN。

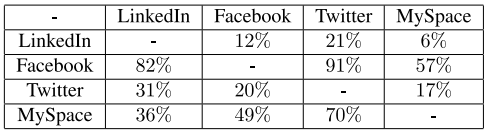
每个OSN中的影响传播对网络的影响是特定的;例如，Twitter和Facebook则完全不同。此外，即使是同一社交网络中的不同级联，也可以用不同的影响传播模型来更好地解释，因此，重叠用户将OSN连接在一起，形成一个OSN的多重结构，由多个OSN通过重叠用户连接在一起，其中每个OSN具有不同的局部传播。本文研究了一类OSNs的多重影响最大化问题，即选择最具影响力的种子节点。

为了演示多路传输与单层传输的区别，作者举了如下例子：



在第一层中是阈值模型，第二层中是IC模型。假设a是种子节点，a会议0.5的概率去影响到b，如果a把b激活了，a和b一起就能把c激活。最后，观察到这两层的激活不能按照IC模型或固定阈值模型合并到单层网络中。

虽然在单一网路使用单一模型的算法已经很普遍了，然而，由于这些算法只考虑单一的影响宣传模型，并不直接适用于MIM，因为每个OSN都有不同的传播模型。如图3所示，重叠用户的比例相当大。



**贡献点**：

1.作者定义了广义确定性子模(GDS)性质，如果每一层都满足该性质，则意味着在多路复用上的整体传播是子模的。

2、给出了一种近似算法KSNs

3。通过实验验证算法的有效性。

**2. MODEL REPRESENTATION AND PROBLEM FORMULATION**

**A influence propagation models**

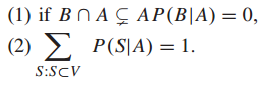
直观上，网络中影响传播模型的思想是清晰的，它是在给定一组种子节点的情况下激活或影响节点的一种方法。

在本文中，由于我们允许一个多路复用器的每一层都有一个不同的影响传播模型，所以我们需要对这个概念有一个技术上的定义

定义1：（model of influence propagation）在图G=(V,E)中一个影响传播模型sigma就是一个函数P，对于每个种子节点A和每个节点S都分配一个概率值



它满足如下两个式子



（1）简单地说，种子节点一定会被激活，（2）确保我们有一个概率分布

表示给定种子集A的预期激活节点数并且有



一个模型是确定性的条件是：对于每个节点A属于V，存在FA使得，确定性意味着在扩散模型中不存在概率，因为最终激活的集合是由种子集唯一确定的。如果sigma是确定的，那么有，我们使用符号来表示，他也是最终集本身。这允许方便地指定集合，例如,σ,τ是确定性模型,T是最终的激活集。

许多的信息传播模型在文献中讨论满足子模性,σ满足，子模性是重要的，因为它保证了一个贪婪的方法的影响最大化问题将有一个近似的比例，现在作者定义了一个比子模块性更强的属性

定义2：KSN

每个σj是一个确定的子模性的模型，pj属于[0,1]，并且有

**引理1**：σ是一个传播模型，如果σ满足GDS，那么σ是子模的。

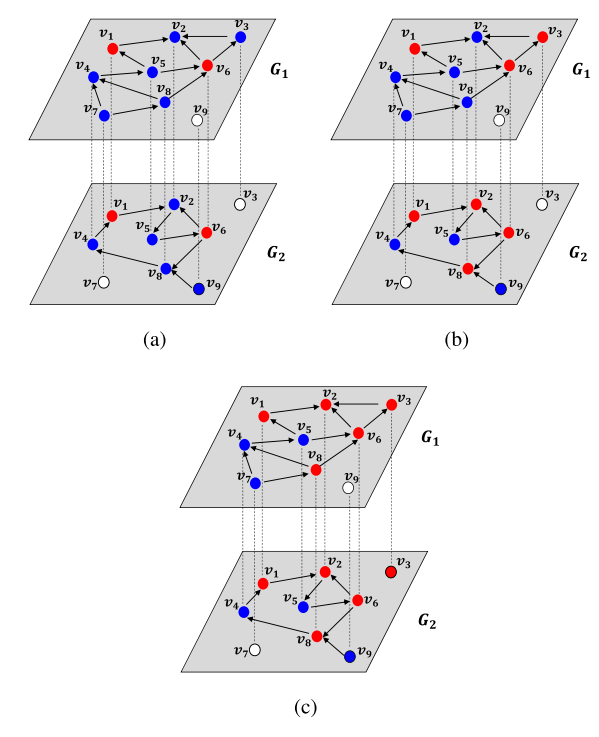
**B notations and multiplex model**

社交网络可以建模为有向图G=(V,E)

定义3：Heterogeneous Multiplex：一个OSNs上的多路是一个列表

，Gi=(Vi,Ei)是一个有向图代表一个OSN和还有在这个图上的一个传播模型σi。如果一个用户属于一个以上的OSN，则在代表该用户的两个节点(每个OSN中有一个节点)之间添加一个层间边缘，这样的用户称为重叠用户; 我们将用O表示重叠用户的集合。作者使用来表示所有用户。

多路复用的影响传播模型σ以下列方式定义，如果一个重叠顶点v在一个图形Gi中被激活，在所有osn中，其相邻层间拷贝被激活;在每个图Gi中的传播根据其传播模型σi进行的。如图4所示：



**C problem definition**

定义4：Heterogeneous Multiplex Influence Maximization：给定一个k层网络

，每层都有自己的传播模型σi，一个正整数l，找一个集合S，S最多包含l个结点，使得被激活的用户数最多，σ（S）最大。

3 APPROXIMATIONS OF MIM

由于单个网络的影响最大化是多重网络影响最大化的一个特例，因此MIM是N p完备的。作者首先证明了影响传播σ是子模的，如果每层的传播满足GDS。为了解决这个问题，作者提出了以下算法

**A greedy approach**

用表示多层网络，并且有传播模型σ，作者已经了σ的子模性，所以，提出的贪心算法就会有的近似比

1）子模性：为了不失去一般性，作者假设每层的结点都是一样的，如果这层没有这个结点，那就加个孤立点。这里计算影响结点数量时，只计算结点的一个副本。

作者将首先考虑一个**简单的情况**:当每个Gi的传播是确定的和子模的。

a）确定的情况：在本节中，让每个Gi中的传播σi是确定性和子模的。给定一组种子集合S，在传播结束后被激活的节点集合可以被表示为，在其中一层网络被激活的节点集合为，可以看出

引理2：让S属于V，可以得到对于所有i，有

引理3：让S，T属于V，对于所有i，有

引理4：有集合S和T属于V，满足：

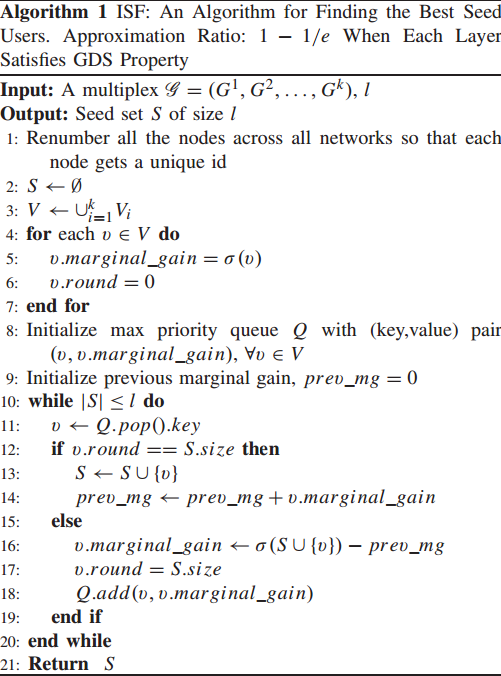


引理5：如果每层的传播模型都是确定的和子模的，那么在多路网络中的传播模型就是子模的。

b）在一定概率情况下：将确定性情形的结果推广到所有网络都满足GDS的情形

定理1：给定一个多路网络g，它有k层，如果每层网络的传播模型符合GDS，那么这个多路网络的传播模型符合GDS

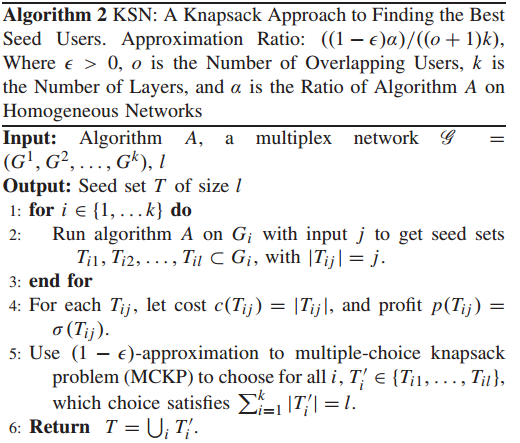
2）一个贪心算法ISF。如下所示：

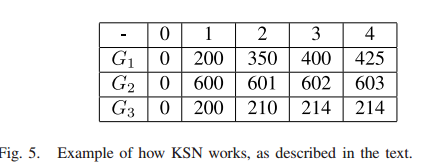


B Parallelizable Multiplex Algorithm

虽然每层的传播模型都符合GDS的话，ISF可以达到的近似比，对于较大的网络规模，ISF的运行时间可能不切实际，所以作者又提出了算法2（KSN）算法，另一种近似算法，根据组件层并行化问题，其难点在于将影响最大化问题的解与分层问题的解相结合，得到MIM的解。KSN通过近似选择背包问题的解来实现这一点。KSN的近似比取决于重叠用户的数量o，层数k。

1）KSN描述：KSN算法将算法A用作输入来计算每层的影响力最大化问题，一个多层网络g，总共有k层，种子的数量为l，对于任意的，算法A同时在每一层运行去找到有j个种子的种子集合Tij，然后，它使用一个近似于背包问题(定义如下)的方法来决定每一层应该播种多少个节点。





第(i, j)个条目激活了在Gi层中播种的j个节点，然后我们使用一个多选题背包算法来选择每一层Gi，要在该层中播种的节点数。

**4. 实验**

在本节中，我们对合成网络和真实网络进行了实验，以证明所提算法的有效性。

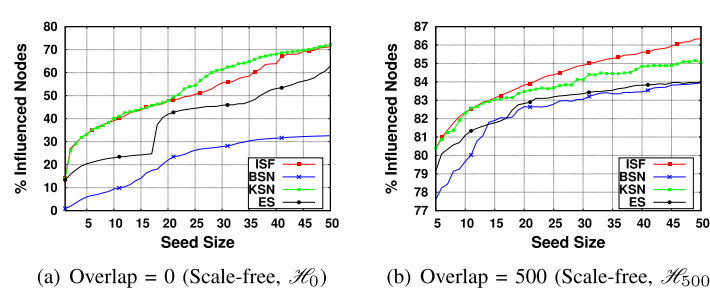
**A Methodology**

ISF，KSN（算法A为CELF++或者IMM），even seed（ES），best single network

**B Synthesized Multiplexes**

作者考虑了一个三层的网络分别为此为此H1,H2,H3，它是由Barabasi–Albert模型生成的，总共有1000个节点和4000条边，平均度数为4。作者给这三层的模型分配为LT，IC和MLT，边权重和节点阈值都是在[0,1]之间随机抽取的。

图6显示了有重叠节点和无重叠节点时各个算法的影响范围



以及在不同网络中各个算法的运行时间

