**CT-IC Continuously activated and Time-restricted Independent Cascade model for viral marketing**

**1.introduction**

由于在线社交网站的快速增长，我们现在体验到，个人的信息和想法通过在线社交网络以极快的速度传播给他人。

虽然提出了几种新的扩散模型，但它们都忽略了影响扩散在实际病毒营销应用中的两个重要方面。首先，在IC和IC- n模型中，当一个节点激活时，它只能激活它的邻居一次，但是在实际中，应该是可以重复激活的（不止一次）。其次，在IC、IC- n和LT模型中，激活过程一直持续到不再发生激活为止，然而，在现实世界中，我们往往有时间限制，因此不能等到影响“完全”传播。

本文提出了一种更为实际的病毒营销应用影响扩散模型，称为连续激活和时间限制的IC(CT-IC)模型。**：**CT-IC模型是IC模型的一种推广，它的不同之处在于:(a)每个活动节点都可以重复地激活它的邻居(b)在给定时间T前处理激活。

**2.related work**

（1）不同的影响传播模型：IC，LT，IC-N考虑了负影响的传播，CLT模型考虑了两个竞争的影响在同一个网络中传播

（2）影响传播模型的参数：在设计上述扩散模型的同时，学习传播概率也很重要。

（3）影响最大化问题及其有效处理方法：定义

影响力最大化问题有两个主要的挑战：首先是np-hard，使用贪心算法，不能在多项式时间内完成估计。

**3.CT-IT模型**

**3.1 motivation**

虽然现有的各种模型都反映了真实的动态影响扩散，但它们忽略了两个主要方面。具体来说，在应用最广泛的影响力扩散模型IC模型中，忽略了时间限制，并且设定每个节点只有一次机会激活其邻居。通过几个例子来说明这两个方面在现实世界中的重要性。

**首先**，每个病毒式营销活动都有时间限制或约束。让我们以苹果的iPhone营销为例。iPhone 4发布后，苹果的营销重点是促进iPhone 4的销售，在积极的市场支持下，大部分人对iPhone 4产生了兴趣，并开始购买。因此，iPhone 4在一段时间内销售一空。然而,苹果并不指望iPhone 4永远引领手机市场。苹果公司无疑开发了另一款尖端手机，iPhone 4S。当苹果公司推出iPhone 4S时，苹果公司显然将广告重点转移到了iPhone 4S上，公众也对这款新产品产生了兴趣。新产品发布后，苹果从未为其老产品做过广告。尽管苹果在iPhone 4S发布后依然在销售iPhone 4

**其次**，在现实生活中存在着反复影响朋友或熟人的机会。让我们考虑另一个例子。假设你买了一个新产品，并在你的Facebook上写了一篇积极的帖子，然后，这个帖子会出现在你的朋友面前，并说服他们对这个新产品有一个积极的看法，这可能会导致他们购买它。重要的是，当你的朋友们稍后再次访问你的主页时，你的朋友可能会被说服购买该产品，尽管他们在发帖时没有被说服。从这个例子中，我们观察到人们通常有多个机会在同一件事情上影响其他人。

现有的模型没有考虑影响扩散过程中的时间约束和连续激活机会，我们提出的CT-IC模型的主要贡献是在影响扩散模型中包含这两个关键方面。

**3.2 模型定义**

该模型是在有向图中建立的。给定一个有向图G=（V，E）每条边上都有一个概率pp0(u,v)代表节点u在某一步激活v的概率，给定给一个种子集合S，和一个时间限制T，CT-IC模型传播方式如下：在t=0时，初始集合中的种子被激活，随后开始进行传播。使用At表示在时间t的活跃种子，在时间t，所有属于At的节点都要以概率试图去激活它的未活跃邻居，其中tu是节点u的激活时间。并且将定义为其中是个单调递减函数并且

fuv的单调递减性质是基于这样一个观察，即说服朋友在每次尝试之后都变得越来越难。作者使用来定义函数f。当所有的激活试验在t时刻结束后，新激活的节点St被包含在激活的节点集中，这时我们有，并且t+1步开始

**3.3 CT-IC模型的性质**

作者用表示在时间t时刻开始时种子集S的影响范围，作者给定种子数量k和时间限制T，在CT-IT模型下的IM问题可以被定义为：找到一个种子集合S，使得S的影响范围最广。

**单调性和子模性**：为了证明系统的单调性和亚模性，我们提出了一个与CT-IC模型等价的易于分析的过程。

IC与CT-IC模型的区别：为了研究在特定情况下CT-IC模型与IC模型的区别，我们引入了一种称为IC与CT-IC的不同率如下。我们给定G=(V,E)，正整数k，和一个时间限制T，定义CT-IC模型的最优解集和IC模型的最优解集为GT和GI我们定义不同率为



dr告诉我们有些时候我们将CT-IC模型看成IC模型也可以获得很好的结果。这个比值可以用来量化IC和CT-IC模型之间的差异

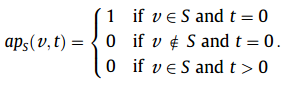
**4. 影响扩散的精确计算**

在这一节中，我们给出了在CT-IC模型下精确的影响评价。虽然，CT-IC模型下的影响传播是np-hard的。然而，如果我们将整个图限制为树状图，它的计算仍然是可处理的，树状图是一种有向图，其中存在从每个节点到根节点的唯一路径。

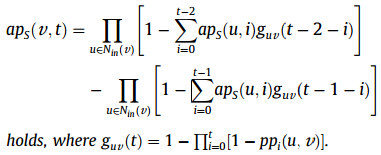
**4.1 树状例子**

假设一个树状图GA=(V,E)，一个种子集合S，一个时间限制T，对于任意的v和时间t，aps(v,t)表示v在时间t被激活的概率，apS,T(v)代表节点v在整个过程结束前被激活的概率。



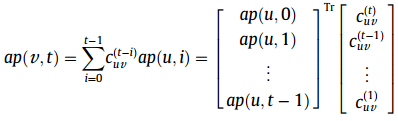


引理2：对于任意的和，有



**4.2 简单路径的例子**

让我们考虑一个简单路径p的影响扩散，它是一个节点序列。对于p上的边(u,v)，节点v在时间t的激活概率为ap(v,t)





**5.影响传播过程算法：**

现有的实验中,我们知道IC模型的贪婪算法在实践中是非常缓慢的，因此，很明显，贪婪算法对于CT-IC模型是绝对不可改变的。我们需要一个对于CT-IC模型可以改变的算法。

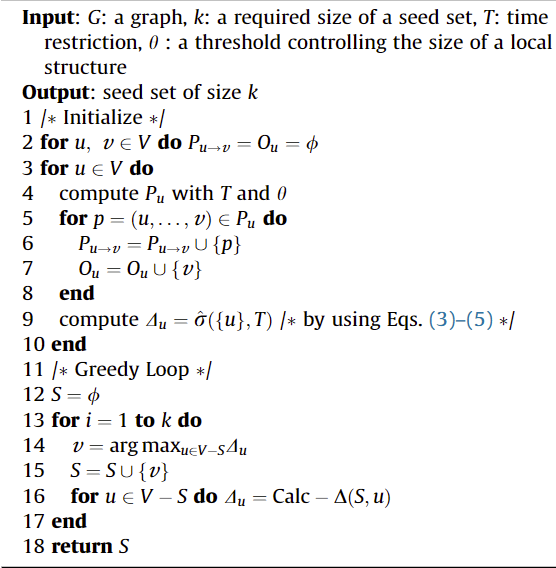
在这一部分中，我们通过扩展IC模型中的IPA算法，提出了连续激活和时间受限的影响路径算法CT-IPA

IPA通过将独立的影响路径作为影响扩散评估的基本单位来评估种子节点的影响扩散。IPA

通过控制影响路径的数量来扩大影响扩散评估，这相当于丢弃传播可能性小于预定阈值的可忽略的影响路径。通过改变影响路径的影响传播定义，可以无缝地完成从IPA到CT-IPA的扩展。在IC模型中，影响路径的影响扩散是通过对路径中每条边的实值传播概率进行多次迭代来获得的。在CT-IC模型中，影响路径的影响传播为infp(,),涉及了一些矩阵乘法。

我们从节点u开始定义重要路径，pu= ，SPu是所有从u开始的简单路径。从节点u到v的重要路径集合为，这代表着u可以从这个集合中的一条路径影响v，最终，节点u的影响区域可以被定义为

CT-IPA算法如下：



**合并多条边**：为了减少算法的运行时间，所以在预处理工作中要把多条变合成一条。在IC模型中，假设从节点u到节点v有好几条边，，这几条边可以合并成一个边e’，它的传播概率为。但是在CT-IC模型中就不是这个样子。

**6.实验**

**6.1实验建立**

数据集：HEP，PHY，EPINION，AMAZON

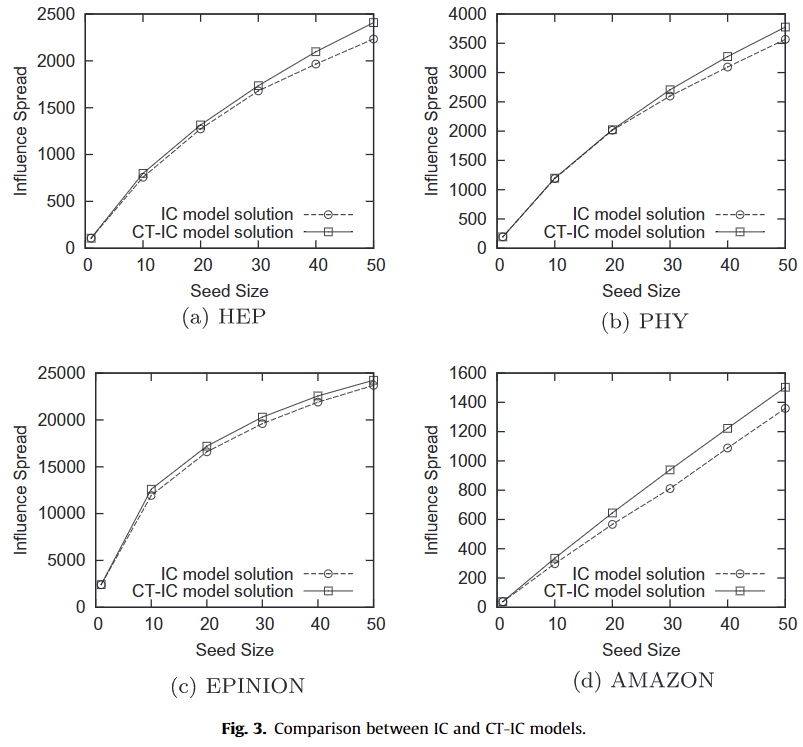
传播概率：1/入度

算法：Random，MaxDegree，Greedy，CT-IPA

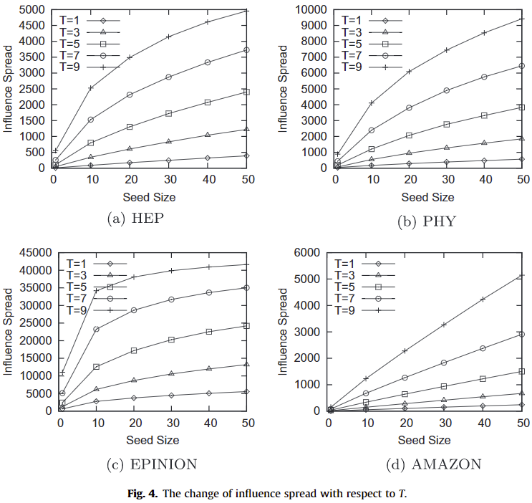
**6.2 CT-IC模型的特征**

CT-IC与IC模型的区别：通过定量和定性两方面的比较，CT-IC比IC模型要更好。我们在这两个模型中分别运行Greedy算法，并且将种子数量设定为从1到50，并且把时间T设定为5。

图3是在两个模型中分别使用greedy算法得到种子集合的影响范围。



改变T得到不同的影响传播，为了找出影响扩散如何随着T的增加而变化，测量T从1到9时的影响扩散，k也从1变化到50。图4说明了影响扩散在四个数据集的结果。在每个数据集上，影响扩散随着时间的增加而增加，这是一个显而易见的结果



**6.3不同算法之间的比较**

下图展示了不同算法之间的影响范围

