**Efficient Distance-Aware Influence Maximization in Geo-social Networks**

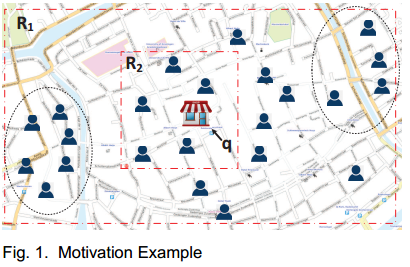
**1.introduction**

随着Web 2.0技术和社交媒体平台的发展，越来越多的公司开始利用社交网络来推广他们的产品。影响最大化是利用社交网络口碑效应的优势，是病毒式营销中的一个关键问题，在文献中得到了广泛的研究。由于影响力的传播是建立在家庭、亲密朋友等之间的信任基础上的，这种营销策略比传统的广告渠道，如电视、报纸等更有效。

在现有的大多数关于影响最大化问题[1]、[2]的著作中，对社交网络中的用户一视同仁，即，每个用户拥有相同的权重。近年来支持地理位置的设备和服务的发展使传统社交网络增加了空间维度成为可能。在进行位置感知的推广时(如在市中心新开的餐厅进行推广)，现有的影响力最大化模型可能由于不了解用户的空间信息而无法满足需求。直观地说，离某个推广地点较近的用户更有可能参加。因此，很自然地认为用户应该根据推广位置得到不同的权重。

在[9]中，每个用户都与多个签入（位置）相关联。基于用户的访问历史数据，推断出给定推广位置的用户之间的传播概率，该问题与本文的研究重点相同。在[8]中，每个用户在二维空间中都有一个位置，给定一个查询区域R，作者试图选择一组只对R中的用户影响最大的用户。然而，这种模式有两个缺点: 1)给定一个推广位置，选择合适的目标区域并不容易。2)忽略了用户与推广地点距离的重要性。

给出了一个例子：如下所示



如例1所示，在演示通过社交网络进行的位置感知促销时，考虑用户与促销地点的距离是至关重要的。在这篇论文中，我们研究了距离感知影响最大化(DAIM)问题[10]结合了两个因素: 影响传播和用户到查询位置的距离。距离感知影响最大化问题的目标是找到一组k个节点，使加权影响最大化。

Challenges：1.结点数量巨大。2. 要满足网上的要求并非易事。对于系统来说，支持在线需求非常重要，因为可能会发出大量的DAIM查询。

Contributions：

1. 我们正式介绍了地理社交网络中的距离感知影响最大化问题

2. 我们提出了MIA- da方法，它扩展了MIA模型以启发式地支持新问题。

3. 提出了一种改进的RIS- da方法，该方法对RIS模型进行了扩展，并对DAIM问题提出了一种无偏估计

4. 我们评估了这些方法在真实地理社交网络上的性能。综合实验验证了该方法的有效性

**2.background**

**2.1**

我们认为地理社会网络是一个有向图G=(V,E)，其中，V代表节点(用户)的集合，E代表边缘(用户之间的关系)的集合，每个v属于V有一个地理坐标（x,y），其中x和y分别表示v的经度和纬度。一个函数w，为每个节点分配对应于二维空间中给定位置q的权重。

传播模型：本篇文章选用IC模型，每条边都有一个独立的传播概率。

**定义1**：距离感知的影响传播：给一个地理社交网络G=(V,E)和一个地理q在二维空间中，节点集合S的距离感知的影响传播为Iq(S),被计算为

其中w(v，q)是v相对于q的权重。

当没有歧义时，我们称距离感知影响传播为影响传播。如第1节中所讨论的，靠近提升位置的用户应该具有更高的受影响优先级。因此，w(v，q)应该与v和q之间的距离成反比。



问题描述：给定地理社交网络G、查询位置q和正整数k，距离感知影响最大化(DAIM)的问题是在G中找到具有最大距离感知影响扩散的k个节点的集合S。



**2.2**

MIA模型和RIS模型是为解决传统的影响最大化问题而设计的。在本文中，我们通过分别扩展这两个模型来开发两种方法。在本节中，我们将介绍这两种模型的细节。

**2.2.1 MIA 模型**

为了解决影响最大化问题，一个重要的步骤是计算一组节点的影响扩散。由于计算影响扩散的难度，MIA模型[3]使用基于树的启发式方法来近似它。MIA模型的思想描述如下。

给定一个社交网络G和两个节点u，v ∈ V，如果两个节点之间有路径，u可以激活V。u,v之间的路径可以表示为，u可以影响v的概率为在MIA模型下，当u和v之间有多条路径时，选择概率最大的路径，因为这条路径为u影响v提供了最大的机会。在u和v之间具有最大概率的路径被称为最大影响路径，表示为MIP(u，v)。然而，许多最大影响路径的概率非常小，因此对于测量影响扩散的贡献是微不足道的。我们使用阈值θ来修剪所有不重要的路径。为了计算集合S影响到v的概率，我们引入MIIA和MIOA

**定义2**：，

MIIA(v)和MIOA(v)是以v为根的树，包括所有可能影响和受g中v影响的节点

**2.2.2 RIS模型**

图分布：概率扩散模型M下的社交网络G可以看作是一个图形分布G={gi}。分布中的每个实例gi对应于通过遵循扩散模型m生成的确定性图.

**定义3**：RR集：给定一个来自G的实例gi和一个节点v，反向可达(RR)集是在gi中可以达到v的节点集。

**定义4：**随机RR集合：RR集是一个随机反向可达集，如果gi是从G中随机抽样的，v是从所有节点v中随机选择的