# ABSTRACT

在本文中，我们引入了一个数学模型，其中捕获了基于流行度的推荐系统的一些突出特征，并尝试利用用户之间的社会关系。**我们表明，在非常普遍的条件下，市场总是趋于稳定的状态**，我们能够给出一个明确的形式。**因此，我们可以准确地说出推荐系统改变市场的程度，并确定用户影响他人的力量。**我们的理论结果与现实世界社交网络的实验相辅相成，表明尽管存在非常有影响力的用户，社交图表可以防止大的市场扭曲。

# INTRODUCTION

推荐系统（RS）是人类与文化领域的算法之间的相互作用的典范。从YouTube视频到亚马逊图书和iTunes或Netflix上的电影，在线选择越来越受到这种算法的介入。RS因其在产品或新闻故事之间产生不平凡关系的能力而受欢迎，因此它们正在成为强大的经济角色。我们打算在本文中探讨RS的这个经济层面。在这方面的一个根本问题，我们打算在本文中解决的问题是：RS可以在多大程度上改变市场？ 例如，假设在线书店开始采用RS。未知书将成为命中，反之亦然？ 典型读者的习惯如何改变？ 大量的工作在个人，用户层面[8,25,26,35]研究了这种问题。

例如，[25]中描述的一个很好的实验表明，即使是基于下载次数提供排名的简单类型的反馈，也可能显着提升市场份额。诸如此类的研究提供了宝贵的见解，但是它们是从他们对于龙润市场的定量准确预测中得出的。此外，完全可以想到，观察到的放大效应是暂时的，或者不能转移到其大小和复杂性与人造环境相比矮小的市场。其他研究试图从实际市场的行为数据中得出结论[11，？，10，23]，但是可用数据的缺乏使得它们成为暂时的。

在本文中，我们尝试采用不同的方法，为由RS管理的市场引入自然模型。我们的模型足够简单，可以对长期行为进行精确的数学分析，同时它捕获了RS的一些相关功能。具体来说，有两种已知在RS中发挥重要作用的成分，我们希望我们的模型代表。**第一个是受欢迎程度，即项目在其市场上的成功程度。一些研究表明，流行度反馈（例如下载次数，用户评分，观看次数等）**可能是在线行为的强大决定因素[25,5,35]。

**第二个因素是，由于缺乏更好的术语，用户之间的“亲戚网络”。**社会关系是这里的范例。他们被认为是形成用户选择的重要因素，并且推荐者尝试以各种方式利用它们[22,31,33,6]。更一般地，RS尝试利用用户之间的不同类型的相似性，例如在线社交网络中的显式链接，消费者选择模式之间的相似性，或者甚至是矩阵分解的复杂计算的结果。所有这些不同的情况可以简单有效地代表，正如我们在本文中所示，假设存在连接用户的（加权）网络。

**简化形式，我们的市场演进数学模型如下。在市场上，我们有通过网络连接的用户（或买家）和产品。当用户购买产品时，它或者遵循其个人倾向（由产品上的私有概率分布引导的模式），或者遵循一定概率的推荐。当它是，它咨询网络中的随机邻居，推荐者节点。推荐者节点从目前为止购买的产品列表中选择随机产品，其概率与当时产品的受欢迎程度成比例，并且所选择的商品由当前买家购买。**

**结果属于非常普遍的条件。首先，用户可以购买的产品以不同的速率。用户u比用户v快十倍，比用户w慢一千倍。第二，跟随推荐的可能性可能会因人而异，每个买方可以自己对建议的个人信任程度，也不同程度地信任他/她的朋友。第三，推荐产品的概率不需要统一，但可以依赖于时间。例如，我们只能推荐最后购买的十件商品，或以其他方式推荐商品推荐的概率。最后，该图可以是任何有向图（所以你可以遵循v的建议，但反之亦然），用户可以在建议机制开始运行之前拥有自己的个人购买历史。这些都不会影响结果 - 市场总是收敛到独特的极限。**

将用户连接到社交网络的网络很自然，但实际上可以用于建模更多的场景。例如，通过假设网络成为团，我们可以使一个项目的流行度与整个市场的销售数量成正比，或者我们可以代表一个影响整个网络的强大的超级用户，而不受它的影响 通过假设网络是一个定向的明星。

我们的结果的一个很好的特点是限制明确地以封闭的形式给出。这使得有可能在分析和计算上探索稳态的重要性质。例如，我们可以很容易地回答我们的激励问题 - 推荐机制能否以重要的方式改变市场？ 我们可以轻松地将每个产品的初始市场份额和稳定状态之间的距离，无论规范如何。我们可以深入了解什么类型的拓扑放大或抑制扭曲效应，我们还可以进行细粒度分析，为每位用户说明影响最终结果的程度。我们分析的副产品是用户影响力的定量观念，确定每个用户在极限时对市场的影响。有趣的是，这个概念证明是个性化PageRank（PPR）的泛化，它与在我们的模型中可以定义的联盟游戏的**Shapley**值有关。

我们的极限定理为现实社会网络的计算探索提供了丰富的景观。**我们探索各种图形拓扑的扭曲效应以及用户可以影响市场的程度。**出现的是一个相当细微的景观，在本文的其余部分全面阐述。也许最有趣的结果是，**真正的社交网络证明自己是一个防止市场扭曲的遏制**，尽管在他们中间有名人的存在。总而言之，本文是**系统调查推荐系统对其运营市场的长期影响的一个方向**。这是一个值得我们关注的重要和迷人的话题。我们希望本文对这个重要问题提供一些可行的见解。

# RELATED WORK

已知产品受欢迎程度对用户行为和选择有强大的反馈效应，并被许多现有的RS所利用。在最有说服力的演示中，[25]表明，在数字文化市场中，仅基于下载次数报告排名显着影响用户选择，并增加市场的不平等和不可预测性。另一个有说服力的例证，[35]证明，当有关物品受欢迎程度的信息时，用户可能会改变主意，扭转以前的选择。Twitter的朋友推荐研究[28]表明，普遍用户的建议比“平均”用户的建议更有可能被接受。几个现有的在线市场使用基于流行度的建议 - 例如Amazon1，Yelp2，TripAdvisor 3，SoundCloud4和Last.fm5都提出了当前趋势的项目。

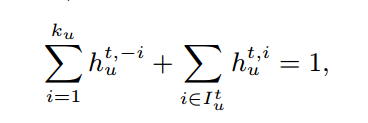
社会关系也被理解为在推荐和采购过程中发挥关键作用，而且越来越多地被现代RS所利用。关于Epinions6网络的研究[7]发现，他/她的朋友也采用了用户采用的大部分项目。同样，对雅虎的研究 脉冲网络[30]表明，用户的兴趣与他/她的社会关系的兴趣高度相关，可以预测，反之亦然。基于社会关系的RS的一个象征性例子是[31]，它使用生成影响模型，用户根据他们和他们的朋友的喜好概率地挑选项目。另一项研究[6]建立了一种推荐算法，假设驱动用户动作的两个力量是潜在的偏好和朋友的影响力。其他一些作品显示或提出建议是在一些社会网络结构上进行的[16,17,33,15,32]

另一方面，对RS的长期市场影响的调查很少; 只要数学模型，现有的工作似乎局限于经济文献。据我们所知，对RS的市场影响的唯一理论研究是[12]。在他们的工作中，单个用户根据他/她的个人口味反复购买两种产品中的一种，或者是迄今为止购买最多的产品的建议。 根据参数的选择，该过程可导致销售集中。他们的模型和我们之间有很多差异 - 特别是我们的模型允许用户互相影响。两个早期的研究，[9]上意见形成共识，达到和[13]，调查用户在一个社会群体互动，个人信仰的融合同样我们如何调查的市场中拥有的社会关系融合。不过，他们的模式在很多方面都远离我们， 例如，我们是随机的，并允许对过去事件的复杂依赖。另一篇论文[14]研究了口碑对产品销售的影响，但着重于均衡策略，以最大化用户满意度。最后，还有一些其他的作品已经分析了RS和市场之间的相互作用，或者是以不同的目标[18,24]或模拟而不是分析[32]

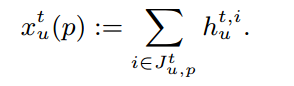
# 3．A MODEL FOR RECOMMENDER SYSTEMS

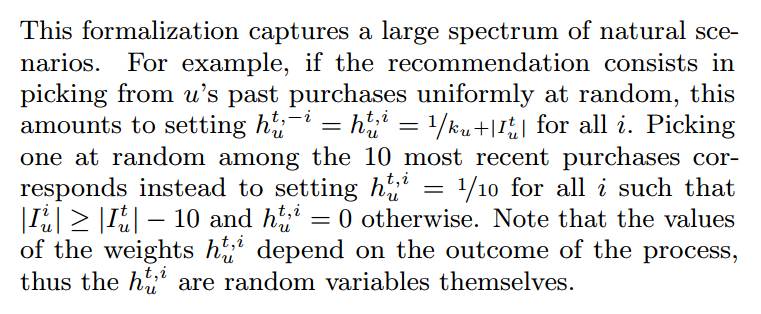
## The weight of history

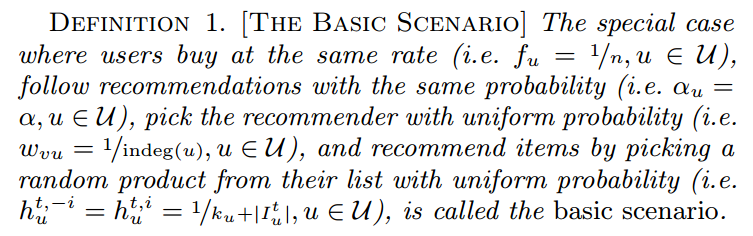
为了得出模型的描述，我们需要定义概率分布，我们用{xtv(p)} p∈P表示推荐者节点的产品被挑中。建立基于人气的建议的**一个合理选择是只需**从v过去购买的多个集合中随机选择一个产品。以这种方式，受欢迎的项目的推荐概率较高。**但是，我们希望通过允许与时间相关的权重来采取更一般的方法。**例如，我们希望能够处理一个情况，即只有十个最近的购买是重要的，或者采购的影响随着时间的推移以某种特定的方式消失，等等。为此，您将重点放在顶点u并回想起我们允许用户在启动系统之前进行ku购买。假设现在我们在时间t> 0，并且Itu表示u以前进行购买的时间集合。我们表示在时间i <t时由u创建的购买的权重为ht,iu。首先，我们假设这些权重被归一化，即所有t> 0



然后，假设在时间点t，u是的推荐节点， 产品p将被节点u推荐并由买方以概率购买







## 备注

我们的模型仅考虑可以多次购买的产品。食物是一个自然的例子，但是有很多文化产品属于这一类：出门到某个俱乐部，参加体育赛事，或者参加特定场地的音乐会只是几个例子。多种购买没有什么意义的商品的情况，如书籍或电影，也是非常有意义的，但将来仍然是未来的工作。值得注意的是，我们的模型可以在一定程度上捕获重复消费，人们往往会重新购买他们最近购买的产品的经验原则[1,2]。更确切地说，我们的模型可以模仿[1]中的一个，用户购买一个新的产品或从他/她的采购历史中的分销中抽取出来的产品。我们可以通过允许图形中的自循环来模拟这个模型：用户可以选择自己作为推荐者节点，并且用正确定义的权重来回购他在过去所购买的产品。

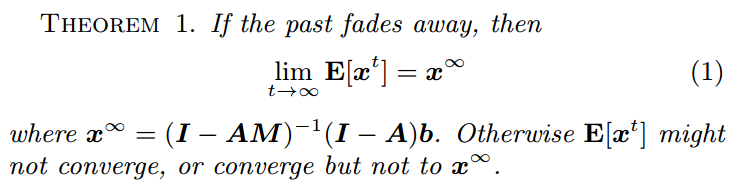
## 3.1 Analysis of the model

符号解释 Recall xt u(p)是在时间t，u是推荐者节点的商品p被推荐和购买的概率。为了简化符号，让xt u(p)简单地表示为xtu。同样，让我们对p的个人偏好，并且让b =（b1，...，bn）。我们想研究xt：=（xt 1，...）的演化。我们定义A = diag，并用M表示G的加权转置邻接矩阵，即Muv = wvu（如果vu 2 / G为0）。对于任意随机向量y =（y1，y2，。），我们用E [y]表示其分量的期望的向量（E [y1]，E [y2]）。

我们现在可以解决我们的主要问题：p的市场份额是否收敛？ 下面的定理表明，如果权重ht,i u满足以下条件，则会发生这种情况。



这种情况非常自然。它表示，过去的购买将最终被遗忘。例如，当推荐产品随机选择均匀或仅最后几次购买具有非零重量时，它是满意的。我们现在可以给出我们的主要结果，这在附录中已被证实.



结束解释标注，我们现在可以解决我们的主要问题：p的市场份额是否收敛？ 下面的定理表明，如果权重ht，i u满足以下条件，则会发生这种情况。过去消失了，即limt！1 E [ht，i u] = 0。这种情况非常自然。

它表示，过去的购买将最终被遗忘。例如，当推荐产品随机选择均匀或仅最后几次购买具有非零权重时，它是满意的。我们现在可以给出我们的主要结果，这在附录中已被证实。

# 4. EXPERIMENTAL RESULTS

在本节中，我们讨论了一系列实验，其中我们的模型的基础图是一个现实世界的社交网络。定理1表示xt在期望收敛到x1。第一组实验表明，期望确实捕获了在极限情况下发生的情况，即方差很小。然后我们调查市场扭曲的问题 - 对于给定的产品p，其市场份额在极限上有多大变化？ 一个相关的问题涉及所谓的影响者。他们的影响是什么？他们可以改变市场多少？ 我们显示，当底层网络是一个真实世界的社交网络市场时，失真可以忽略不计，甚至“名人”也不会产生重大影响。反之亦然，还有其他现实的情况，例如作为“媒体大亨”的超级节点的存在，由于推荐机制，某些节点可以征服市场的重要份额。我们还给出了一个经验法则，根据他们的2跳邻居来确定顶级影响者。最后，我们探讨我们的理论结果未涵盖的推荐者，其中推荐使用超线性概率的产品。我们发现结果从图形变化很大，而且在很大程度上取决于产品的起始份额。现在让我们来描述实验设置。

## 4.1 Experimental setup

我们在六个以下真实世界图上运行我们的实验

Google+ Twitter SNAP Twitter LAW Slashdot Yelp Facebook