# Hierarchical Community-Level Information Diffusion Modeling in Social Networks

近来，在线社交网络正日益成为社交互动的热门平台。了解信息如何在这样的网络中传播对于社交搜索中的个性化和推荐非常重要。

在本文中，我们提出了一个社区层次信息传播（HCID）模型来捕捉社交网络中的信息传播过程。我们通过引入用户的话题受欢迎度的概念使我们的模型能够描述既是话题感知（信息涉及哪个话题）又是信息源感知（信息来源）的信息传播过程。我们提出社区层次性的概念，而不是假设社交社区的同质性，跨层次社区间的信息传播是从高层到低层是单向的。

我们设计了Gibbs抽样算法来推导模型参数，并提出了两种信息扩散预测任务的预测方法，即转发预测和级联预测。比较实验在两个真实数据集上进行。结果表明，与现有工作相比，我们的模型取得了实质性的进步。

# Leveraging Cross-Network Information for Graph Sparsification in Influence Maximization

当我们在解决大规模的Influence maximization问题时，一个有效的策略是使用图形稀疏化作为预处理步骤，通过去除一部分边使原始网络变得更简洁和易于处理。在这项工作中，提出了一种Cross-Network Graph Sparsification（CNGS）模型，以利用在源网络中预先检测到的影响力中心知识来预测和去除对目标网络中影响传播作出很少贡献的边。实验结果表明，提出的CNGS模型进行图的稀疏化可以在IM的效率和有效性之间取得良好的折中，即现有的IM贪心算法可以更有效地运行，而影响传播的损失尽可能小在稀疏的目标网络中。

# From Footprint to Friendship: Modeling User Followership in Mobile Social Networks from Check-in Data

在本文中，我们旨在解决移动社交网络（MSN）中两个关键因素之间的关联：用户之间的社交关系网络和用户的空间移动模式。具体而言，我们研究用户空间分布对MSN社交关系形成的影响。基于MSN用户的地理定位数据（签到记录）和社会关系数据，我们提出了一个模型，称为neighborhood-cardinality-based model（NCBM），将用户的空间流动性的homepoint/hotpot特点和长尾社会关系度分布都考虑在内。我们为每个用户定义一个基本数量，即所谓的neighborhood cardinality，来衡量其他MSN用户访问他的附近区域的频率和频率。 NCBM的一个核心原则是：用户v跟随用户u的概率服从用户u的neighborhood cardinality的幂律分布。我们的模型是在两个大的check-in数据集上进行评估的：Brightkite和Gowalla。我们的实验结果表明，所提出的形成模型为捕捉MSN用户的移动模式和社交关系之间的相关性提供了有用的范例。

# Predictive Network Representation Learning for Link Prediction

在本文中，我们提出了一个predictive network representation learning（PNRL）模型来解决结构链接预测问题。所提出的模型定义了两个学习目标，即观察结构保存和隐藏链接预测。为了将这两个目标集成在一个统一的模型中，我们制定了一个有效的抽样策略来选择给定网络中的某些边缘作为假设的隐藏链路，并将剩余的网络结构视为在训练模型时观察到的。通过对这两个目标进行联合优化，不仅可以提高节点表示的预测能力，而且还可以在表示空间中学习额外的链接预测知识。在四个真实世界的数据集上的实验证明了所提出的模型优于其他流行和最先进的方法的优越性。

# Top-K Influential Nodes in Social Networks: A Game Perspective

影响力最大化是病毒式营销的基础，目的是在特定的传播模型下寻找最大化影响扩散的最top-k种子节点。在本文中，我们从博弈论角度研究影响最大化。我们提出了一个协调博弈模型，在这个模型中，每个人都根据与其网络邻居协调的好处做出决定，。我们的模型作为一些现有模型的推广，比如多数表决模型和线性阈值模型。在广义模型下，研究贪婪算法的影响最大化的难度和逼近保证。我们还结合了几种策略来加快算法运行速度。实验结果表明，加速后的算法明显优于其他的启发式，比原始贪婪算法快三个数量级。

# Predicting Which Topics You Will Join in the Future on Social Media

社交媒体用户每天都会在几乎所有有可能的话题上发送数百万条微博。如果我们可以预测未来用户将加入哪些主题，那么很容易确定哪些主题会变得流行以及主题可能吸引什么类型的用户。对于许多应用程序来说，这也是非常有意义的。在这项研究中，我们调查预测用户是否会根据其发布历史来加入主题的问题。我们引入了一个新的深层卷积神经网络与外部神经记忆和注意机制来解决这个问题。用户的发布历史和主题是用外部神经记忆体系结构建模的。基于卷积神经网络的匹配方法被用来构建用户和主题之间的关系。最终的决定是基于这些匹配的结果。为了训练和评估所提出的方法，我们从Twitter收集了一个大规模的数据集。实验结果表明，所提出的方法可以显着优于其他方法。与最先进的深度神经网络相比，我们的方法在F1分数和MAP10分别达到了18.2％和28.9％。

# Computational Social Indicators: A Case Study of Chinese University Ranking

许多专业组织定期发布社会指标报告来监测社会进步。尽管它们有合理的结果和社会价值，早期在社会指标计算方面存在着三个问题：1）劳动密集型数据收集问题，2）数据不足问题，3）依赖专家的数据融合。为此，我们通过探索丰富的多渠道Web数据，提出了一种新的基于图结构的多渠道社会指标计算排序方案。对于每个通道，该方案分别以简单图和超图来呈现半结构化和非结构化数据。然后根据它们之间的相关性把这些信道分成不同的簇。之后，使用统一的模型来学习聚类智能空间，分别对每个空间进行排序，并将这些排序进行融合以产生最终的排名。我们以中国的大学排名作为案例研究，并通过真实世界的数据集来验证我们的方案。值得强调的是，我们的方案适用于计算其他社会指标，如教育程度。