

计算机科学：大数据挖掘与社交网络可视化分析

1: 引言

1. 本篇文章的背景是什么？

全球最大的社交网络平台——Facebook 社交平台

2. 本篇文章的知识空缺？

网络的 Assortativity coefficient（同配系数）

同配性（Assortativity），用作考察度值相近的节点是否倾向于相互连接。在社交网络中，节点倾向于与度数相近的节点相连。

如果总体上度大的节点倾向于与度大的节点相连，那么该网络的度是正相关的。或者称网络是同配的。如果度大的节点倾向于与度小的节点相连，那么该网络的度是负相关的，或者称网络是异配的。

原文链接：https://blog.csdn.net/sword_csdn/article/details/120273195

2: 引言

1. 本篇文章的研究问题是什么？

Facebook 社交网络的结构特性(structure)

2. 本篇文章的研究目的是什么？

以数量分析(Quantitative Analysis)方法探讨 Facebook 社交网络的特性，与传统的网络特性比较并总结其基本特性比如 homophily, clustering, the small-world effect, heterogeneous distributions of friends, and community structure.

3: 文献综述

1. 简要谈谈在文献综述部分中文章间的逻辑是什么？他们之间有什么联系吗？

描述了过往对社交网络的相关研究，联系体现在方法上的相似性（量化分析方法），平台的对比性（与 Twitter,Yahoo 等其他平台的对比）。

4: 方法论

1. 简要谈谈本篇文章用的什么研究方法？

- (1) 数学建模方法，把社交网络对应到数学研究中的“图模型”并构建相关模型和其约束
- (2) 可视化图表分析的方法：绘制了各类数量分析的变化图
- (3) 对比实验方法：通过对比实验得到科学可靠的结论

5: 方法论

1. 简要谈谈本篇文章是怎样收集数据分析的？对此你有什么收获？

从图的下面几个性质进行了可视化的数据分析：Degree distribution、Path lengths、Component sizes、Clustering coefficient and degeneracy.

收获是要得到可靠的结论需要全方位多角度的数据分析。

6: 结果

1. 简要谈谈本篇文章是如何报告相关结果的？

相关结果主要集中在“Result”部分，在最后 Discussion 和一开始的 Abstract 部分也有简要提及。

相关结果主要是通过数据的量化分析和统计得到的。

7: 讨论

1. 本篇文章是如何解释研究结果的？

从社交图的几个重点部分的分析：Degree distribution、Path lengths、Component sizes、Clustering coefficient and degeneracy.，结合量化统计结果，并用可视化图表的方式结合了 common sense 得到相关结果

8: 讨论

1. 本篇文章的研究局限性是什么？

- (1) 没有做更细致的数据建模并进行特征工程
- (2) 没有将相关模型数据代码开源
- (3) “A pure power-law was seen to be inappropriate for the degree distribution of Facebook, although hubs certainly exist”(原文)

9: 结论

1. 本篇文章对相关研究领域的贡献是什么？

最新的最全面的正对 facebook 社交网络的结构化特性化分析研究

10: 结论

1. 本篇文章是否给出了相关建议？如果有是哪些建议？

有，作者建议可以进一步构建相关图算法和社交网络分析方法。

2. 通过对本篇文章的阅读，你最大的收获是什么？

文章发现社交圈最大的局限就是大部分都是同国家的同龄人，所以需要拓宽到更多国家多年龄段的朋友拓宽自己的社交局限。