统计物理思想的大数据与社交网络研究获进展

作者：writer 来源：中国科学院理论物理研究所

本文原地址：https://www.iikx.com/news/progress/1173.html

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

近年来线上社交网络(online social network)的规模不断扩大。比如微信和Facebook等的用户已经超过了十亿量级。在如此巨大的网络中如何精确并快速地量化某一用户在整个网络中的影响力，从而识别最有影响力的个体或群体，成为一个具有挑战性的问题。

目前为止，大部分现有算法的时间复杂度至少为O(N)。也就是说，这些算法所花的时间随着网络规模的增大而迅速地增加。 最近，中国科学院理论物理研究所副研究员金瑜亮(共同一作)及其合作者(中山大学副教授胡延庆、西南交通大学博士生纪圣堃、新加坡高性能计算所研究员冯凌、美国波士顿大学教授Gene Stanley、以色列巴依兰大学教授Shlomo Havlin)在国际期刊《美国国家科学院院刊》(PNAS)上发表了题为Local structure can identify and quantify influential global spreaders in large scale social networks的研究论文。该论文提出了一个称为PBGA的新算法，其理论时间复杂度与网络规模无关，从而解决了以上难题。

PBGA算法的提出受到了物理学中临界现象的启发：早在2002年，Newman就提出网络中的信息传播过程可以对应到一个经典的物理学问题——渗流相变(percolation transition)。渗流相变是一个标准的临界相变。对应于临界相变中的关联长度，该研究提出了传播半径的概念。基于网络中每个节点在传播半径范围内的局域网络结构信息，可以精确地度量该节点的传播能力。传播半径只与距离临界点的距离有关，而与网络规模无关。 在微博、Facebook、QQ、Twitter等实际网络上的测试结果表明(如图)，PBGA算法的时间复杂度确实和网络规模基本无关。

基于简单外推估算，对于全局的Facebook网络，PBGA算法比经典贪心算法(NGA)将快约1010倍。该算法不仅高效，而且克服了在规模较大的网络上无法得到完整的全局信息的困难，在病毒式营销(viral marketing)等电子商务领域有重要应用前景。

此项目得到中科院率先行动百人计划的资助。(来源：中国科学院理论物理研究所)

更多 科学进展 请访问 https://www.iikx.com/news/progress/

本文版权归原作者所有，请勿用于商业用途，爱科学iikx.com转发