

尊敬的吴老师，您好

我最近看了几篇关于网络信息传播模型的论文，重点阅读了这篇论文：“Learning Social Network Embeddings for Predicting Information Diffusion”，这篇论文提出的CDK(Content diffusion Kernel)模型简洁明了，我想先从这篇论文开始。下面是我对这篇论文的一些理解。

CDK模型的创新之处在于大胆地抛弃了网络的拓扑结构，转而从热传导的现象来对信息传播建模。具体的：我们先将网络中的节点映射到N维欧氏空间中，假设某时刻节点A转发了某条消息，那么我们想象在节点A处亮起一簇火苗，随着时间的流逝，节点A附近的节点温度不断升高，一旦某节点(假设为B)的温度达到某个阈值，那么B被激活。在给定的时间内，按时间先后依次激活的节点序列则构成一条转发链。

CDK模型的目标则是根据训练样本，来确定网络节点的坐标。当然，我们并不需要具体的去模拟这一物理过程来训练模型，训练过程可以简洁地转化为一个排序过程。

相较于传统的IC模型，CDK的计算速度很快，大概分别是minutes与days的量级。在没有网络拓扑结构的情况下，CDK的预测精度能与IC模型旗鼓相当，甚至更好。

作者开源了论文代码，基于Torch7框架实现。我从头到尾读了一遍。

老师您觉得这篇论文可以做下去吗？

学生王超民，2016年9月25日