

《程序员》2012年第6期

个性化推荐的十大挑战

周 涛

电子科技大学，互联网科学中心
百分点推荐技术研究所

<http://blog.sciencenet.cn/home.php?mod=space&uid=3075&do=blog&id=554630>

一、数据稀疏性问题

- MovieLens-4.5%
- Netflix-1.2%
- Bibsonomy-0.35%
- Delicious-0.046%
- 高阶关联
- 缺省打分
- 迭代寻优
- 转移相似性

二、冷启动问题

- 如何针对新用户新产品进行推荐
- 利用属性/文本/标签信息
- 利用品类信息
- 利用跨领域的交叉信息
- 团购和维品会的例子

三、大数据处理与增量计算问题

- 用户商品连接数目已经达到了百亿至十亿规模时代
- 复杂性低：时间、空间、代码
- 可并行化程度高
- 局部近似增量算法
- 自适应算法

四、多样性与精确性的两难困境

- 精确性高的算法不一定就是好的算法
- 信息价值
- 凹透镜和凸透镜
- 个性化的精髓
- 如何评价多样性
- 直接调整推荐列表
- 物质扩散和热传导的混合算法

五、推荐系统的脆弱性问题

- 噪音
- 恶意用户
- 道高一尺，魔高一丈
- 信誉系统

六、用户行为模式的挖掘和利用

- 新用户倾向于选择热门的商品，而老用户对于小众商品关注更多
- 新用户所选择的商品相似度更高，老用户所选择的商品多样性较高
- 不同用户打分的模式也很不一样
- 用户针对不同商品的行为模式也不一样
- 对用户移动模式的理解对于基于位置的推荐至关重要

七、推荐系统效果评估

- 四大类：准确度、多样性、新颖性、覆盖率
- 准确度又可以分为四大类：预测评分准确度、预测评分关联、分类准确度、排序准确度
- 分类准确度包含**6**种常见指标：准确率、召回率、准确率提高率、召回率提高率、**F1**指标和**AUC**值
- 第一层次：数据和算法指标
- 第二层次：商业应用的关键指标
- 第三层次：用户体验

八、用户界面与用户体验

- 推荐结果的可解释性，对于用户体验有至关重要的影响
- 协同过滤有明显的优势，譬如亚马逊基于商品的协同过滤在发送推荐的电子邮件时会告诉用户之所以向其推荐某书，是因为用户以前购买过某些书。相对地，矩阵分解或者集成学习算法就很难向用户解释推荐结果的起源。
- 用户更喜欢来自自己朋友的推荐而不是系统的推荐
- 建立一个可以进行A/B测试的系统

九、多维数据的交叉利用

- 交叠社会网络
- 相当比例的用户都具有交叉购物的习惯
- 交叉推荐的重要商业价值
- 交叉推荐可以提高多样性
- 交叉推荐可以解决冷启动问题

十、社会推荐

- 用户更喜欢来自朋友的推荐而不是被系统“算出来的推荐”
- 社会影响力被认为比历史行为的相似性更加重要
- 通过社会关系的分析，可以大幅度提高从科研文献到网购商品推荐的精确度
- 社会推荐有两方面的效果：一是增加销售，二是在销售后提高用户的评价
- 如何利用社会关系提高推荐的精确度？
- 如何建立更好的机制以促进社会推荐？
- 如何将社会信任关系引入到推荐系统中？

推荐参考的综述和专著

- 刘建国, 周涛, 汪秉宏, 个性化推荐系统的研究进展, 自然科学进展 19 (2009) 1-15.
- 刘建国, 周涛, 郭强, 汪秉宏, “个性化推荐系统评价方法综述”, 复杂系统与复杂性科学, 2009年第6卷3期, 1-10页.
- 朱郁筱, 吕琳媛, "推荐系统评价指标综述", 电子科技大学学报 41 (2012) 163-175.
- 项亮, 陈义, 王益, 推荐系统实践, 图灵出版社, 2012.
- 苏萌, 柏林森, 周涛, 个性化: 商业的未来, 机械工业出版社, 2012.
- F. Ricci, L. Rokach, B. Shapira, P. B. Kantor, Recommender Systems Handbook: A Complete Guide for Scientists and Practitioners, Springer, 2011.
- D. Jannach, M. Zanker, A. Felfernig, G. Friedrich. Recommender Systems: An Introduction. Cambridge University Press, 2011.
- G. Adomavicius, A. Tuzhilin, Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions, IEEE Transactions on Knowledge and Data Engineering 17 (2005) 734-749.
- L. Lü, M. Medo, C. H. Yeung, Y.-C. Zhang, Z.-K. Zhang, T. Zhou, Recommender Systems, Physics Reports, <http://dx.doi.org/10.1016/j.physrep.2012.02.006>