



推荐算法综述

纪幼纯

2017.3.15



研究背景及意义

随着网络技术和信息技术的飞速发展，网络上的信息爆炸性增长，人们经常陷入“信息迷航”的问题中。当他们浏览网页或者新闻，往往很难描述自己想要什么，在网上浏览一圈之后，还是找不到自己感兴趣或者需要的信息。

因此，推荐系统应运而生。作为一个桥梁，个性化推荐技术将用户和信息资源联系在一起。一方面，推荐系统可以帮助用户过滤掉不需要的信息，另一方面，还可以将一些不热门但是用户真正需要的信息展示给用户，这就很好的缓解了“信息过载”和“信息迷航”问题[1]。

[1] Dong, Y., S. Liu, and J. Chai. *Research of hybrid collaborative filtering algorithm based on news recommendation. in Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), International Congress on.* 2016. IEEE.



最早开始研究推荐系统的是明尼苏达大学，提出了最早的协同过滤推荐技术，后来，斯坦福大学的**Marko Balabanovic**在美国人工智能协会上推出个性化推荐系统[2]。之后，**Yahoo!**和**Google**推出了个性化首页服务，亚马逊将个性化推荐运用到电子商务中，分析用户购买行为，使用协同过滤推荐技术为用户推荐商品。

[2]Balabanovic, M., *Learning to surf: multiagent systems for adaptive web page recommendation*. 1998.



编号	名字	年份	使用的推荐方法	应用领域
1	IMDb	1990	协同过滤	影音
2	亚马逊	1994	基于物品的协同过滤	商品
3	潘多拉电台	2000	基于内容的推荐	音乐
4	last.fm	2000	协同过滤	音乐
5	Netflix	2006	混合推荐	视频
6	今日头条	2012	混合推荐	新闻
7	网易云音乐	2013	混合推荐	音乐

[3]Arekar, T., M.R. Sonar, and N. Uke, *A Survey on Recommendation System*. 2014.



David W. McDonald等人（**2000**）[4]提出了一个基于专家知识定位的推荐系统，在推荐系统中加入专家知识模块，适合专业性较强的推荐。**US Manikrao**等人（**2005**）[5]提出基于语义匹配动态地选择网络服务，运用语言匹配来为用户搜索进行推荐。**Moon-Hee Park**等人（**2007**）[6]提出移动端的推荐算法，首先用贝叶斯网络进行用户偏好建模，然后基于上下文，如地理位置、时间、天气等信息为用户进行推荐服务。**Frank E. Walter**等人（**2008**）[7]提出了一种基于信任模型的社交网络推荐。**S Debnath**（**2008**）[8]等人也提出一种基于社交网络和基于内容特征相似度的推荐算法。

-
- [4] McDonald, D.W. and M.S. Ackerman. *Expertise recommender: a flexible recommendation system and architecture*. in *Proceedings of the 2000 ACM conference on Computer supported cooperative work*. 2000. ACM.
- [5] Manikrao, U.S. and T. Prabhakar. *Dynamic selection of web services with recommendation system*. in *Next Generation Web Services Practices, 2005. NWeSP 2005. International Conference on*. 2005. IEEE.
- [6] Park, M.-H., J.-H. Hong, and S.-B. Cho. *Location-based recommendation system using bayesian user's preference model in mobile devices*. in *International Conference on Ubiquitous Intelligence and Computing*. 2007. Springer.
- [7] Walter, F.E., S. Battiston, and F. Schweitzer, *A model of a trust-based recommendation system on a social network*. *Autonomous Agents and Multi-Agent Systems*, 2008. 16(1): p. 57-74.
- [8] Debnath, S., N. Ganguly, and P. Mitra. *Feature weighting in content based recommendation system using social network analysis*. in *Proceedings of the 17th international conference on World Wide Web*. 2008. ACM.

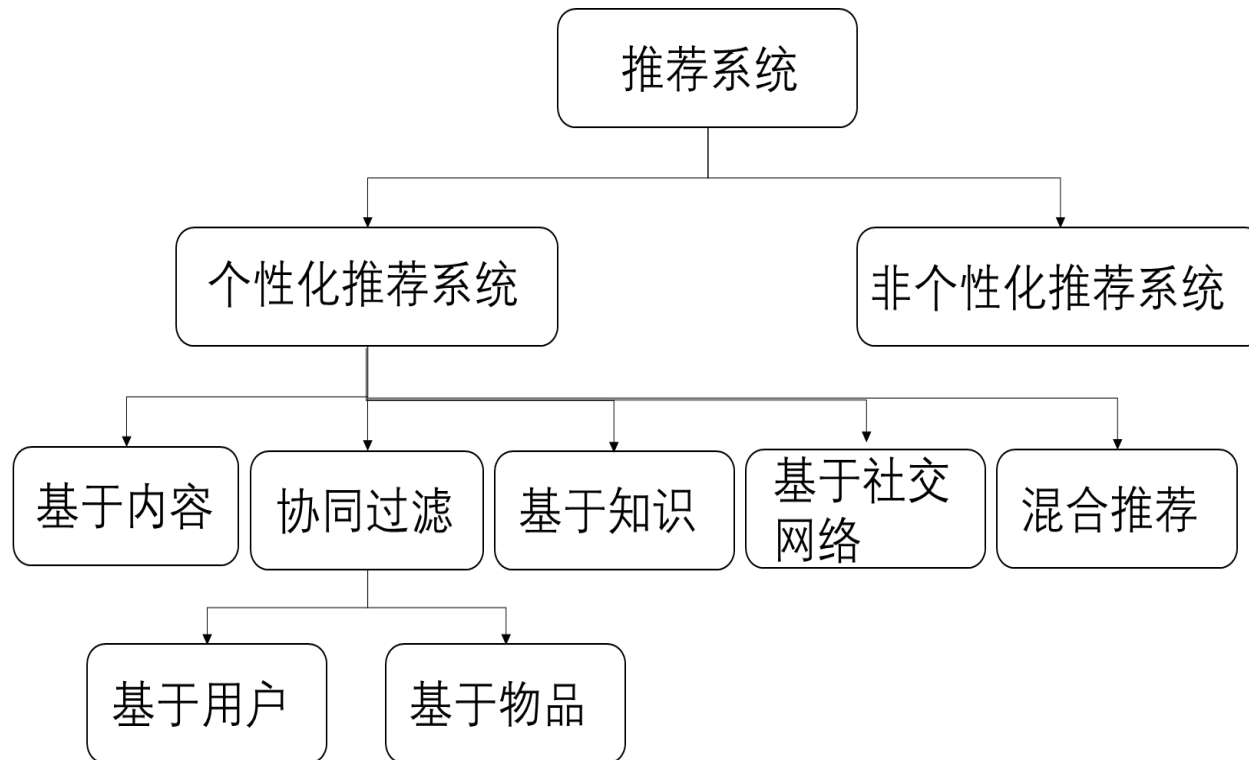


James Davidson等人（**2010**）[9]分析了**Youtube**视频的推荐方法，除了基于视频相似度给用户推荐视频，还会根据用户在网站上的行为为用户进行建模，综合视频相似度和用户模型为用户推荐感兴趣的视频。 **P**

Covington等人（**2016**）[10]针对**Youtube**视频更新快、用户规模庞大的特征，使用深度学习来构建候选模型和排序模型，为用户进行推荐。

[9] Davidson, J., et al. *The YouTube video recommendation system*. in *Proceedings of the fourth ACM conference on Recommender systems*. 2010. ACM.

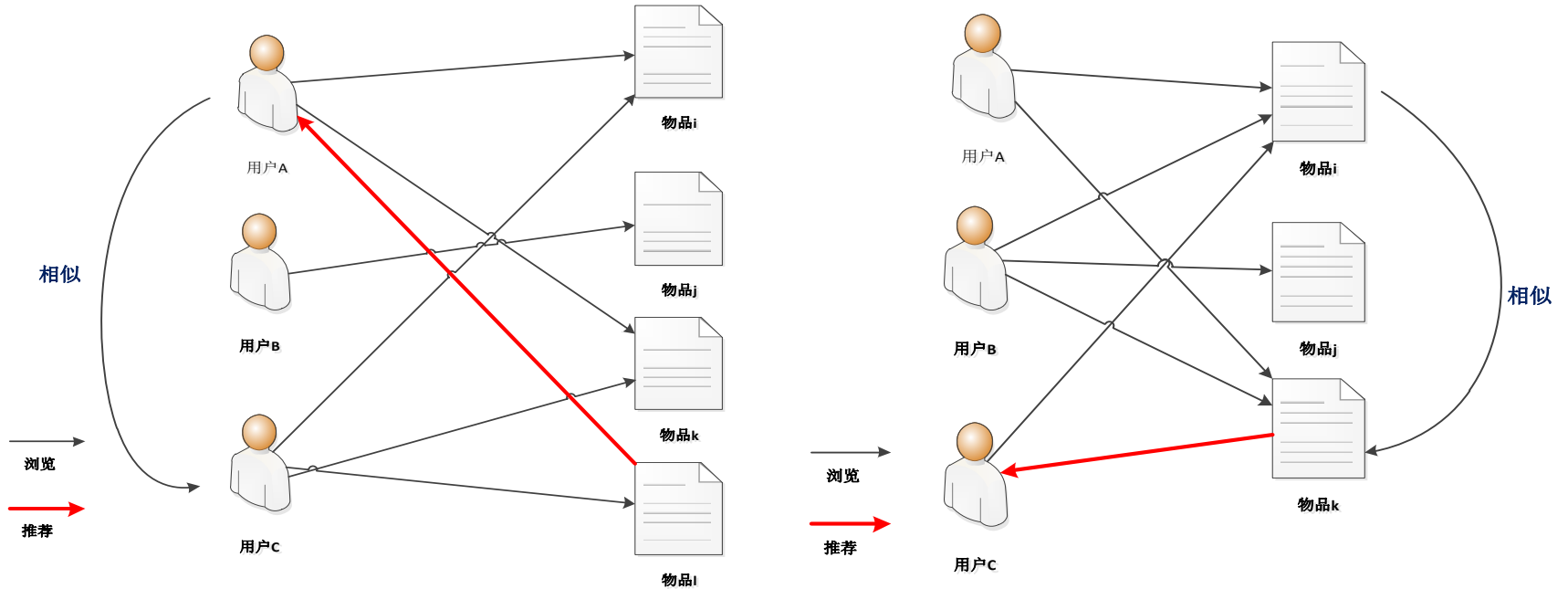
[10] Covington, P., J. Adams, and E. Sargin. *Deep neural networks for youtube recommendations*. in *Proceedings of the 10th ACM Conference on Recommender Systems*. 2016. ACM.



[11] Arekar, T., M.R. Sonar, and N. Uke, *A Survey on Recommendation System*. 2014.



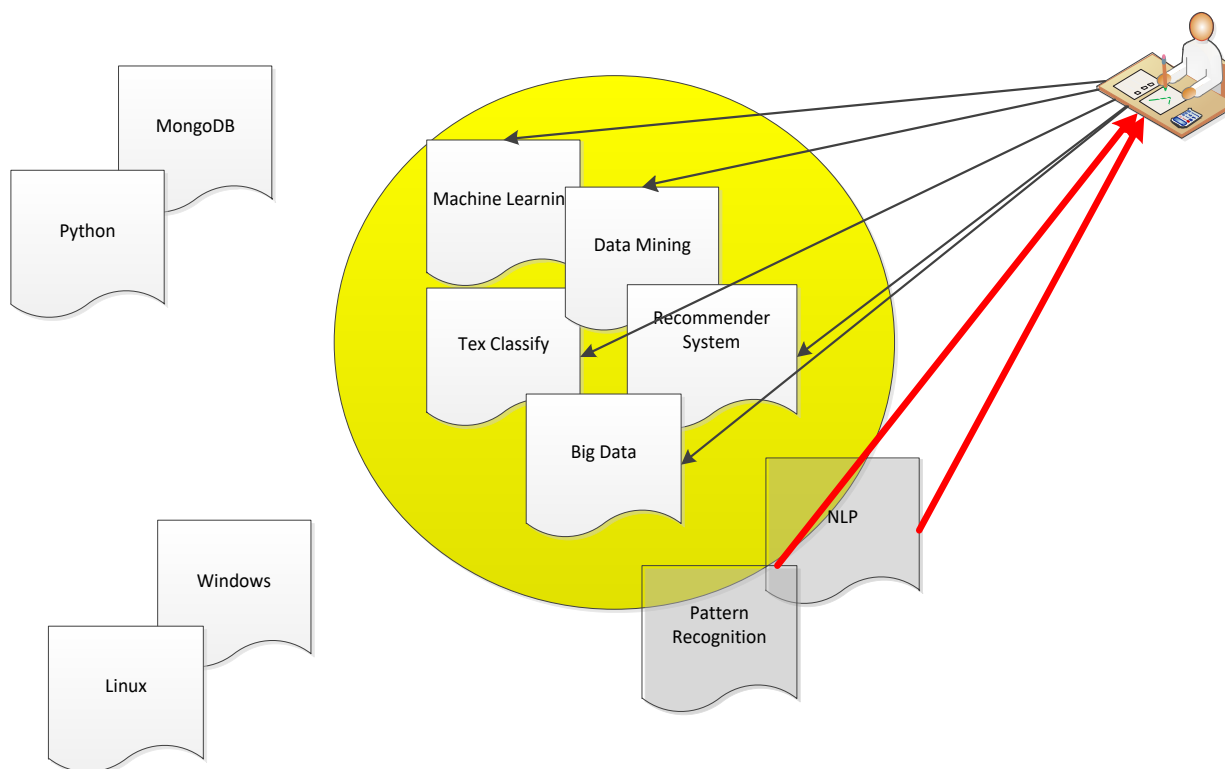
协同过滤推荐算法



[12] Breese, J.S., D. Heckerman, and C. Kadie. *Empirical analysis of predictive algorithms for collaborative filtering*. in *Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence*. 1998. Morgan Kaufmann Publishers Inc.



基于内容的推荐算法



[13] 任磊, 推荐系统关键技术研究. 2012, 华东师范大学.



基于知识的推荐算法

基于知识的推荐算法在传统的推荐算法中，加入了专家知识库，适用于特定领域的推荐，推荐准确度也会较高，基于知识的推荐算法也可以看成基于内容的推荐算法的扩展。

基于知识的推荐引入专家知识库，可以将商品特征做语义扩展，将语义相似的内容进行推荐。同时，基于知识的推荐算法可以识别特定的命名实体、主体、新词等，有利于提高推荐的准确度。但是基于知识的推荐需要构建相应的专家知识库，及时更新知识库，所以人工成本较高。

[14] 张龢, 周年喜, and 张英, *基于语义扩展的个性化知识推荐技术研究*. 情报理论与实践, 2009(08): p. 101-104.



基于社交网络的推荐算法

随着近年来**Twitter**、新浪微博、微信等社交网络的兴起，基于社交网络的推荐算法也得到越来越广泛的应用。基于社交网络的推荐算法是在传统的推荐算法中加入用户的社交网络信息，再根据用户社交网络中的朋友之间的相似度，将朋友感兴趣的物品推荐给用户，可以将其归为协同过滤推荐算法的扩展。

传统的协同过滤推荐算法只利用用户对物品的评分矩阵，人与人之间的兴趣相似度也只是基于相似的评价模式，引入基于社交网络的推荐算法的原因，一方面是社交网络中人与人之间的兴趣相似度更高，二是用户之间有信任度，推荐可解释性较强。

[15] 张富国, *基于社交网络的个性化推荐技术*. 小型微型计算机系统, 2014(07): p. 1470-1476.



混合推荐算法

混合推荐算法	
组合方式	说明
加权	加权组合多种推荐算法
变换	根据问题背景和实际情况或要求决定变换采用不同的推荐算法
混合	同时采用多种推荐算法给出多种推荐结果给用户选择
层叠	先使用一种推荐算法产生粗糙的推荐结果，再使用第二种推荐算法在此基础上做出更加精准的推荐
特征扩充	一种算法产生附加的信息特征嵌入到另一种推荐算法的特征输入中
元级别	用一种推荐算法产生的模型作为另一种推荐算法的输入

[16] Jannach, D., et al., *Recommender systems: an introduction*. 2010: Cambridge University Press.



不同推荐算法的优缺点对比

推荐算法的优缺点对比		
推荐算法	优点	缺点
基于内容的推荐算法	推荐结果直观、可解释性强； 无冷启动和稀疏性问题； 能为特殊兴趣爱好的用户进行推荐	新用户问题； 有限的信息进行分析； 推荐结果单一
基于协同过滤的推荐算法	容易发现用户潜在兴趣，不需要领域知识； 随着时间推移性能提高； 推荐结果多样性高	新用户问题； 数据稀疏性问题； 可扩展性问题； 初始推荐质量较差
基于知识的推荐算法	扩展物品的属性； 提高推荐精度	专家知识库更新较慢； 需要人工干预
基于社交网络的推荐算法	引入社交网络信息； 推荐可解释性强	用户隐私问题； 社交网络信息获取难度大



个性化推荐算法在新闻领域中的研究与应用

新闻的阅读方式随着互联网的进步而改变，由传统的新闻消费模式——通过实体报纸订阅模式变成了由成千上万的新闻媒体网站获取新闻。像谷歌新闻和雅虎新闻，从各种新闻来源获取新闻，并提供来自世界各地新闻的总体视图。但是对新闻服务网站而言，最大的问题是新闻的数量对用户而言太过庞大，所以，帮助用户找到他们可能感兴趣的新闻至关重要[17]。

以算法分发新闻是“互联网+”对传统新闻行业的颠覆，受到了越来越多人的重视。“今日头条”作为移动端第二大新闻应用，主要就是利用用户行为数据来分析用户的兴趣模型，并实现以算法来分发新闻。此外，学术界也提出了许多新闻推荐的方法。

[17] 中国互联网络信息中心, 2016 年中国互联网络新闻市场研究报告 2017.1.



M Medo等人（2009年）[18]提出了一种自适应的新闻推荐模型，因为新闻热度的提升，用户评分会有相似趋势，所以本文模拟新闻的传播方式，根据新闻受欢迎程度给用户推荐新闻。J Liu等人（2010年）[19]基于用户访问日志的分析，构建一个贝叶斯框架，基于部分用户的行为和人们点击后新闻的趋势，预测用户对当前新闻的兴趣程度。并使用基于内容和基于协同过滤推荐算法混合地算法来推荐新闻。他们将推荐系统部署在谷歌新闻网上，数据表明混合推荐算法能提高推荐地准确度并增加用户访问量。

[18] Medo, M., Y.-C. Zhang, and T. Zhou, *Adaptive model for recommendation of news*. EPL (Europhysics Letters), 2009. 88(3): p. 38005.

[19] Liu, J., P. Dolan, and E.R. Pedersen. *Personalized news recommendation based on click behavior*. in *Proceedings of the 15th international conference on Intelligent user interfaces*. 2010. ACM.



L Li等人（2011年）[20]提出一个可扩展地两阶段个性化新闻推荐系统。第一阶段对新闻进行聚类，考虑新闻的特有属性，如新闻内容、访问模式、命名实体、热门程度等，第二阶段考虑用户的兴趣特征，将新闻的选择框架和用户兴趣结合起来，为用户推荐具有新颖性和多样性的新闻。W IJntema等人（2010年）[21]提出了基于本体和TF-IDF的新闻推荐技术，基于语义相似度给用户推荐新闻。

[20] Li, L., et al. *Scene: a scalable two-stage personalized news recommendation system*. in *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*. 2011. ACM.

[21] IJntema, W., et al. *Ontology-based news recommendation*. in *Proceedings of the 2010 EDBT/ICDT Workshops*. 2010. ACM. - 16 -



Thanks