

第七章、推荐系统实例

讲师：武永亮



教学目标

- 了解外围架构
- 理解推荐系统架构

目录

1

外围架构

2

推荐系统架构

3

推荐引擎的架构

外围架构

- 推荐系统和网站的其他系统的接口。目前流行的推荐系统界面共性：
 - 通过一定方式展示物品，主要包括物品的标题、缩略图和介绍等。
 - 很多推荐界面都提供了推荐理由，理由可以增加用户对推荐结果的信任度。
 - 推荐界面还需要提供一些按钮让用户对推荐结果进行反馈，这样才能让推荐算法不断改善用户的个性化推荐体验。

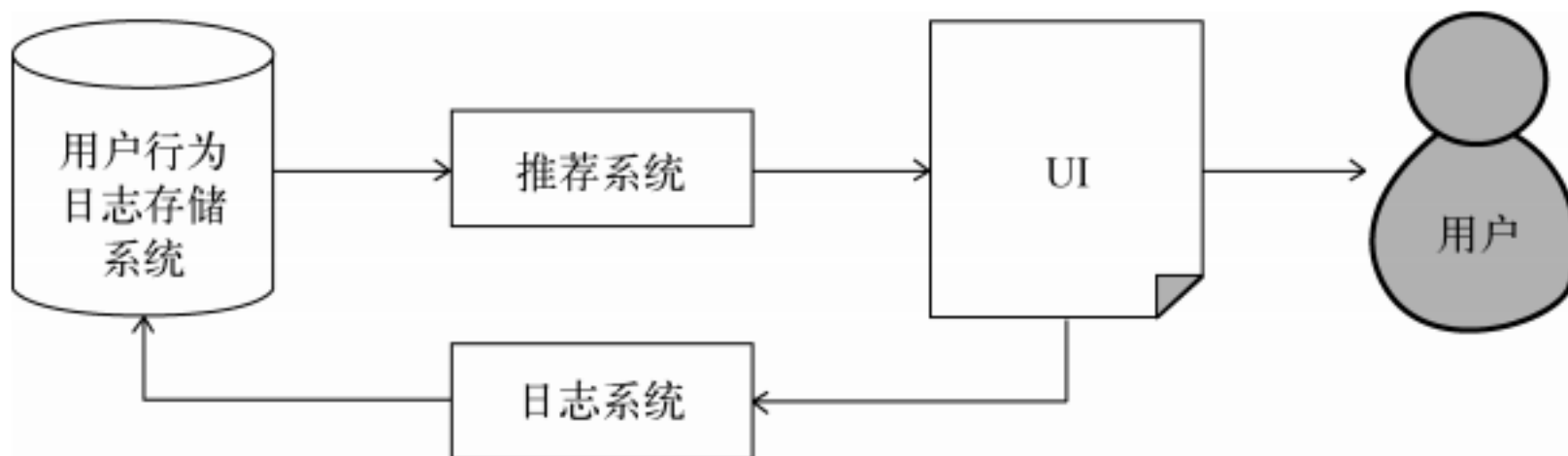


图7-1 推荐系统和其他系统之间的关系

外围架构

- 个性化推荐算法依赖于用户行为数据，而在任何一个网站中都存在着各种各样的用户行为数据。那么如何存取这些数据就是推荐系统需要解决的首要问题。

表 7-1 电子商务网站中的典型行为

行 为	用户类型	规 模	实时存取
浏览网页	注册/匿名	大	×
将商品加入购物车	注册	中	√
购买商品	注册	中	√
收藏商品	注册	中	√
评论商品	注册	小	√
给商品评分	注册	小	√
搜索商品	注册/匿名	大	×
点击搜索结果	注册/匿名	大	×
分享商品	注册	小	√

目录

1

外围架构

2

推荐系统架构

3

推荐引擎的架构

推荐系统架构

- 推荐系统是联系用户和物品的媒介，而推荐系统联系用户和物品的方式主要有3种。如果将这3种方式都抽象一下就可以发现，如果认为用户喜欢的物品也是一种用户特征，或者和用户兴趣相似的其他用户也是一种用户特征，那么用户就和物品通过特征相联系。

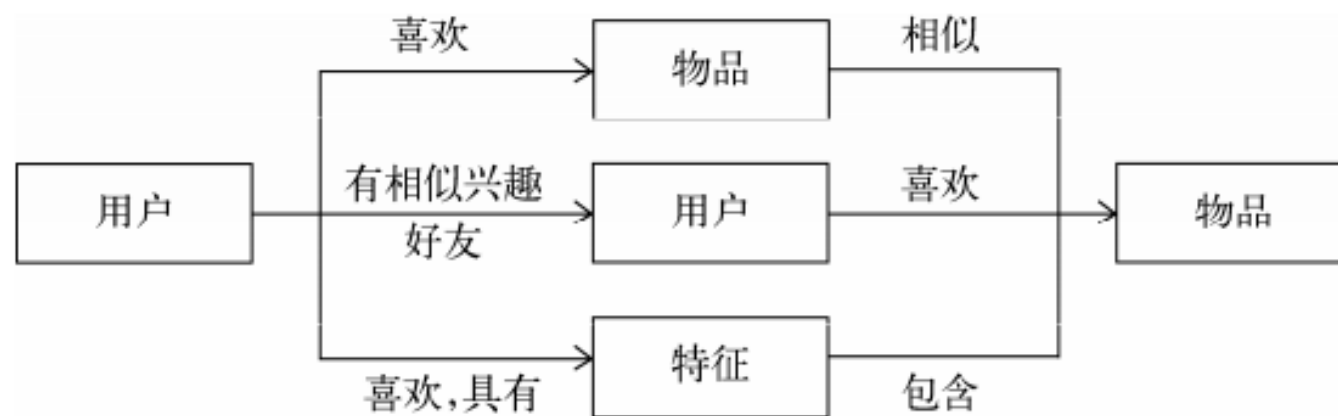


图7-2 3种联系用户和物品的推荐系统

推荐系统架构

- 根据上面的抽象，可以设计一种基于特征的推荐系统架构。

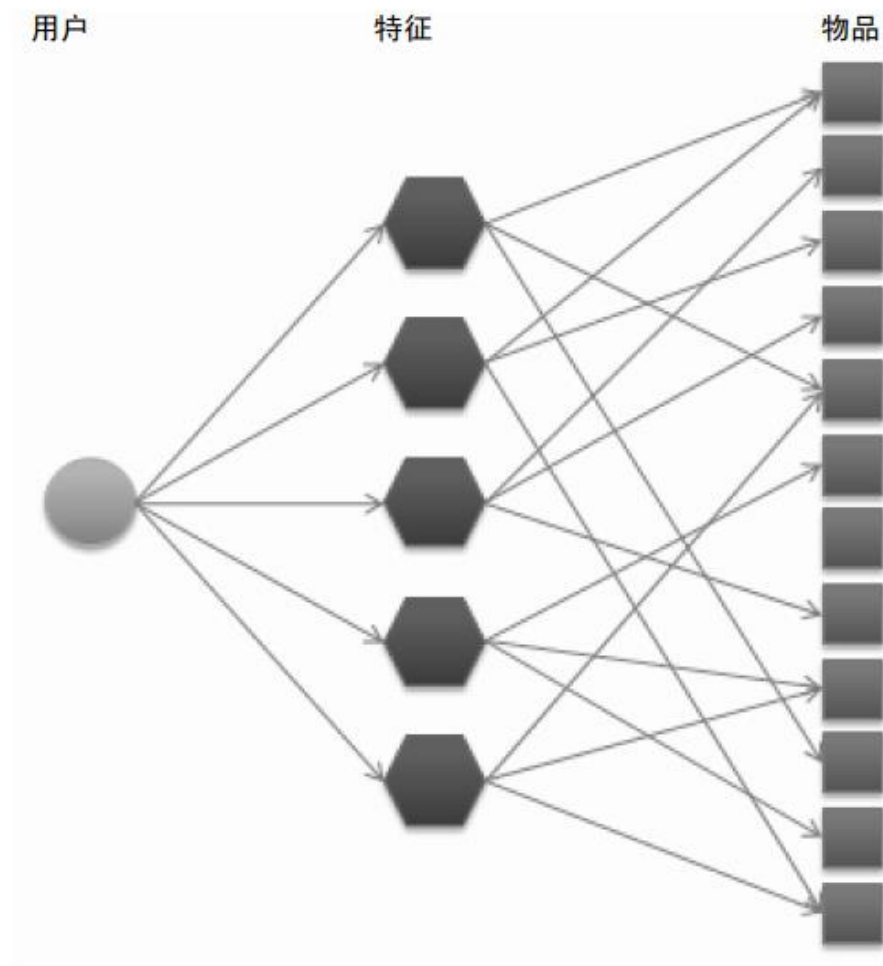


图7-3 基于特征的推荐系统架构

推荐系统架构

- 用户的特征种类非常多，主要包括如下几类：
 - 人口统计学特征
 - 用户的行为特征
 - 用户的话题特征

推荐系统架构

- 推荐系统的推荐任务也有很多种
 - 将最新加入的物品推荐给用户。
 - 将商业上需要宣传的物品推荐给用户。
 - 给用户推荐不同种类的物品。
 - 给用户混合推荐。
 - 对于不同的产品推荐不同新颖度的物品。
 - 考虑到用户访问推荐系统的上下文。

推荐系统架构

- 推荐系统的推荐任务也有很多种

Related to Items You've Viewed

You viewed

Customers who viewed this also viewed



Apple iPhone 4 Black
Smartphone 32GB
~~\$569.00~~



Apple iPhone 4 16GB
Smartphone Black
iPhone
~~\$699.99~~ **\$529.00**



Apple iPhone 4S 16GB
(White)
~~\$649.95~~



Apple iPhone 4S 16GB -
AT&T - Black
~~\$539.00~~



Toshiba Portege R835-P01
13.3-Inch...
~~\$899.99~~ **\$749.88**

[View or edit your browsing history](#)

Recommendations for You in Books



Data Mining: Practical
Machine...
Ian H. Witten, Eibe Frank,
Mark A. Hall
Paperback
~~\$69.95~~ **\$38.49**
[Fix this recommendation](#)



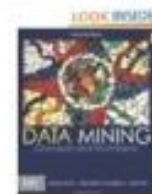
Pattern Recognition and
Machine Learning
Christopher M. Bishop
Hardcover
~~\$94.95~~ **\$60.56**
[Fix this recommendation](#)



Machine Learning: An
Algorithmic...
Stephen Marsland
Hardcover
~~\$72.95~~ **\$59.28**
[Fix this recommendation](#)



Speech and Language
Processing
Daniel Jurafsky, James H.
Martin
Hardcover
~~\$149.99~~ **\$110.26**
[Fix this recommendation](#)



Data Mining: Concepts and
Techniques...
Jiawei Han, Micheline
Kamber, Jian Pei
Hardcover
~~\$74.95~~ **\$48.99**
[Fix this recommendation](#)

[See more recommendations](#)

推荐系统架构

- 推荐系统的推荐任务也有很多种

Popular Among Your Friends on Facebook

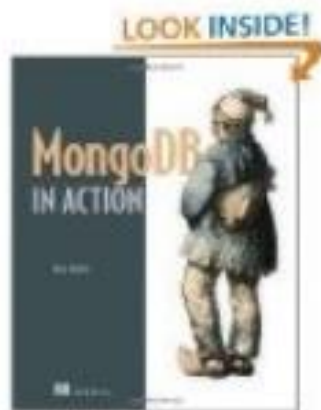


Product	Price	Rating	Friends
The Beatles 1 ~ The Beatles	\$32.49	★★★★☆ (1,195)	7 friends like this:
Gone with the Wind, 75th An... by Margaret Mitchell	\$10.98	★★★★☆ (370)	4 friends like this:
Inception (Two-Disc Ed... Blu-ray ~ Leonardo DiC...	\$12.99	★★★★☆ (970)	6 friends like this:
The Dark Side Of The Moon - Immersi... ~ Pink Floyd	\$79.24	★★★★☆ (107)	6 friends like this:

推荐系统架构

- 推荐系统的推荐任务也有很多种

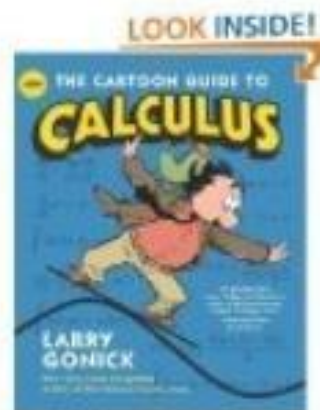
New For You®



MongoDB in Action
(Paperback) by Kyle Banker

★★★★☆ (3) \$29.69

[Fix this recommendation](#)



The Cartoon Guide to Calculus (C...
(Paperback) by Larry Gonick

★★★★☆ (3) \$12.39

[Fix this recommendation](#)

推荐系统架构

- 推荐系统的推荐任务也有很多种

李宗盛 Jonathan

出生日期: 1958-07-19

国家/地区: 台湾

厂牌: 滚石唱片

流派: 流行

风格: 流行/民谣

官方网站: <http://blog.sina.com.cn/lizongsheng>

豆瓣FM 正在播放「李宗盛」

douban.fm/?context=channel:0|musician_id:104294

推荐系统架构

- 如果要在一个系统中把上面提到的各种特征和任务都统筹考虑，那么系统将会非常复杂，而且很难通过配置文件方便地配置不同特征和任务的权重。因此，推荐系统需要由多个推荐引擎组成，每个推荐引擎负责一类特征和一种任务，而推荐系统的任务只是将推荐引擎的结果按照一定权重或者优先级合并、排序然后返回。

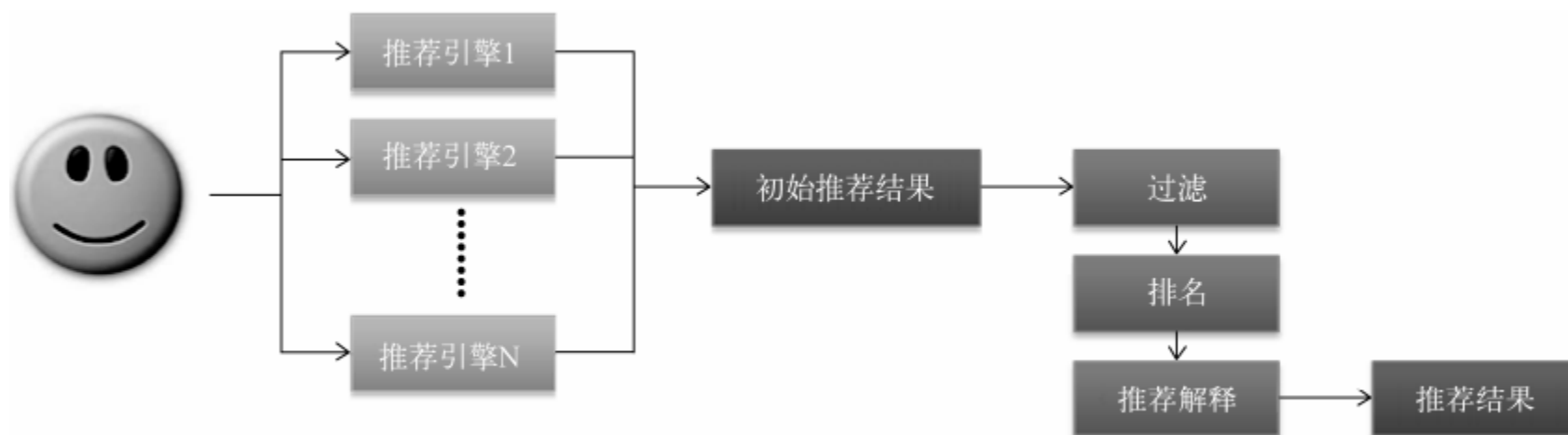


图7-8 推荐系统的架构图

推荐系统架构

- 这样做还有两个好处。
 - 可以方便地增加/删除引擎，控制不同引擎对推荐结果的影响。对于绝大多数需求，只需要通过不同的引擎组合实现。
 - 可以实现推荐引擎级别的用户反馈。每一个推荐引擎其实代表了一种推荐策略，而不同的用户可能喜欢不同的推荐策略。有些用户可能喜欢利用他的年龄性别作出的推荐，有些用户可能比较喜欢看到新加入的和他兴趣相关的视频，有些用户喜欢比较新颖的推荐，有些用户喜欢专注于一个邻域的推荐，有些用户喜欢多样的推荐。我们可以将每一种策略都设计成一个推荐引擎，然后通过分析用户对推荐结果的反馈了解用户比较喜欢哪些引擎推荐出来的结果，从而对不同的用户给出不同的引擎组合权重。

目录

1

外围架构

2

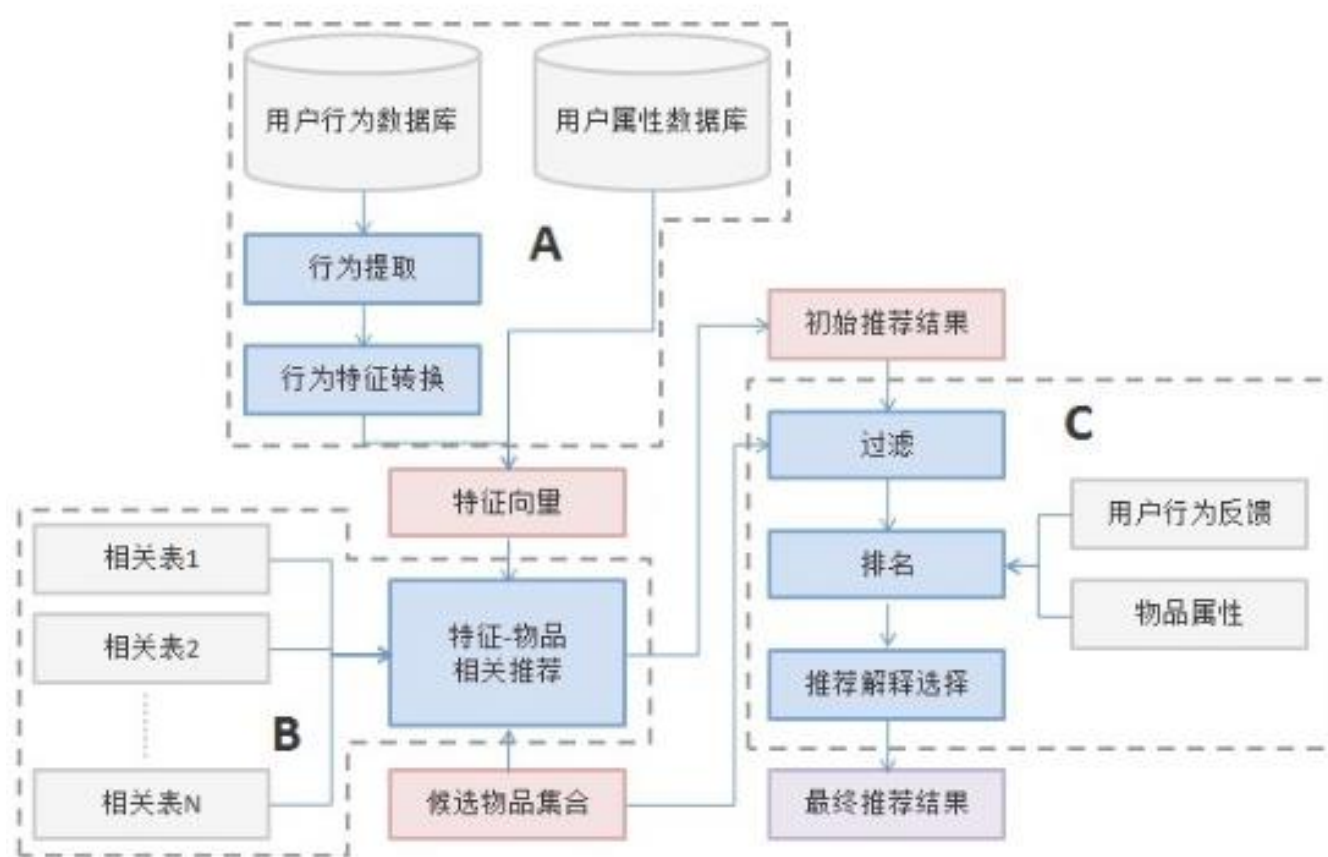
推荐系统架构

3

推荐引擎的架构

推荐引擎的架构

- 推荐引擎使用一种或几种用户特征，按照一种推荐策略生成一种类型物品的推荐列表，推荐引擎架构主要包括3部分。



推荐引擎的架构

- 用户的特征包括两种，
 - 一种是用户的注册信息中可以提取出来的，主要包括用户的人口统计学特征。
 - 另一种特征主要是从用户的行为中计算出来的。
- 一个特征向量由特征以及特征的权重组成，特征向量相关因素：
 - 用户行为的种类
 - 用户行为产生的时间
 - 用户行为的次数
 - 物品的热门程度

推荐引擎的架构

- 在得到用户的特征向量后，我们可以根据离线的相关表得到初始的物品推荐列表。

表7-2 离线相关表在MySQL中的存储格式

src_id	dst_id	weight
特征ID	物品ID	权重

推荐引擎的架构

- 在得到初步的推荐列表后，还不能把这个列表展现给用户，首先需要按照产品需求对结果进行过滤，过滤掉那些不符合要求的物品。一般来说，过滤模块会过滤掉以下物品：
 - 用户已经产生过行为物品
 - 候选物品以外的物品
 - 某些质量很差的物品

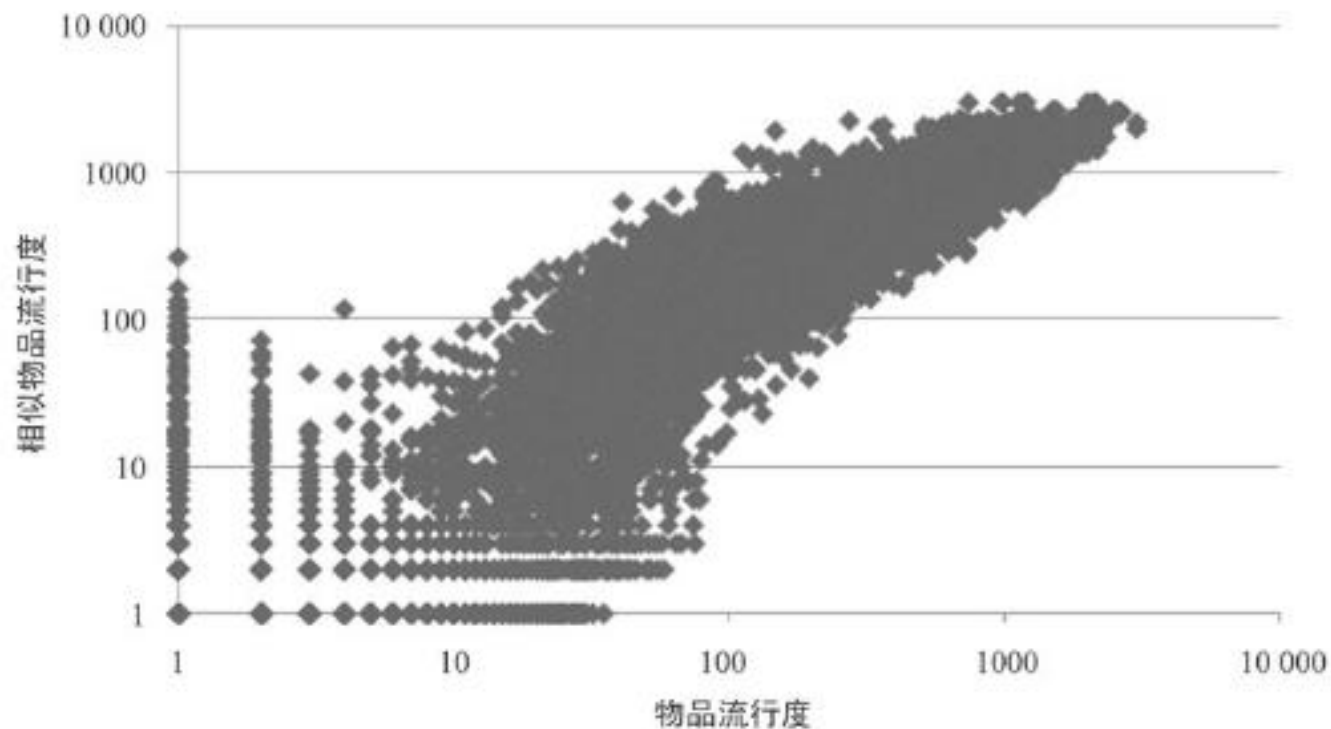
推荐引擎的架构

- 经过过滤后的推荐结果直接展示给用户一般也没有问题，但如果对它们进行一些排名，则可以更好地提升用户满意度，一般排名模块需要包括很多不同的子模块。
- 新颖性排名目的是给用户尽量推荐他们不知道的、长尾中的物品。虽然前面的过滤模块已经过滤掉了用户曾经有过行为的物品，保证了一定程度的新颖性，但是用户在当前网站对某个物品没有行为并不代表用户不知道这个物品。

$$P_{ui} = \frac{P_{ui}}{\log(1 + \alpha \cdot \text{popularity}(i))}$$

推荐引擎的架构

- 多样性也是推荐系统的重要指标之一。增加多样性可以让推荐结果覆盖尽可能多的用户兴趣。当然，这里需要指出的是提高多样性并不是时时刻刻都很好。



推荐引擎的架构

- 提高多样性的方法:
 - 第一种提高多样性的方法是将推荐结果按照某种物品的内容属性分成几类，然后在每个类中都选择该类中排名最高的物品组合成最终的推荐列表。
 - 第二种提高推荐结果多样性的方法是控制不同推荐结果的推荐理由出现的次数。

推荐引擎的架构

- 排名模块最重要的部分就是用户反馈模块。用户反馈模块主要通过分析用户之前和推荐结果的交互日志，预测用户会对什么样的推荐结果比较感兴趣。
- 如果推荐系统的目标是提高用户对推荐结果的点击率，那么可以利用点击模型（click model）预测用户是否会点击推荐结果。点击预测的主要问题是预测用户看到某个推荐结果时是否会点击。那么要进行点击率预测，首先需要提取特征。在推荐系统的点击率预测中可以用如下特征预测用户 u 会不会点击物品 i ：
 - 用户 u 相关的特征；
 - 物品 i 相关的特征；
 - 物品 i 在推荐列表中的位置；
 - 用户之前是否点击过和推荐物品 i 具有同样推荐解释的其他推荐结果；
 - 用户之前是否点击过和推荐物品 i 来自同样推荐引擎的其他推荐结果。

内容回顾

1

外围架构

2

推荐系统架构

3

推荐引擎的架构

Thank you !