



[www.qconferences.com](http://www.qconferences.com)

[www.qconbeijing.com](http://www.qconbeijing.com)

[www.qconshanghai.com](http://www.qconshanghai.com)

伦敦 | 北京 | 东京 | 纽约 | 圣保罗 | 上海 | 旧金山

London · Beijing · Tokyo · New York · Sao Paulo · Shanghai · San Francisco



QCon全球软件开发大会

International Software Development Conference



@InfoQ



infoqchina

软件  
正在改变世界!

# 个性化的天猫

张奇（得福）天猫推荐算法团队

2014·04·20

天猫 TMALL.COM





张奇

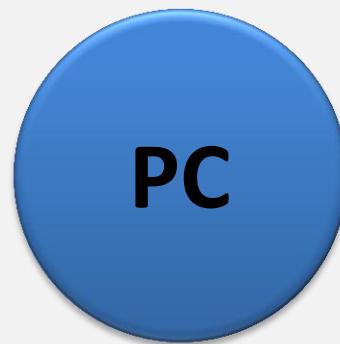
中国科学技术大学，本科（2005），博士（2010），方向：信息检索  
2010年-2012年，阿里云搜索与广告团队  
2012年-现在，天猫推荐算法团队(10个人、15@7月)

天猫推荐业务介绍/推荐系统架构历程

双11个性化会场案例

阿里巴巴大数据竞赛/天猫推荐算法大赛





**成交 : 22%**

去年同期 : 15%

**成交 : 6%**

去年同期 : 0%

看了又看

看了又看

大牌街

大家好

迪

Ja W

11

此商品已  
猜你喜欢

品  
品  
品  
品  
品

品牌店铺

Casarte/卡萨帝  
BCD-346WSCV 三  
门/346L/电脑控温/  
¥ 6988

BCD-230STCE 冰箱  
¥ 2690

看了又看

春夏男鞋特卖折扣包邮  
大牌正品 轻便舒适

跑鞋之王全场3折起  
沐春出行 活力运动

迷情夏 畅销精选套装  
撞色拼接 风景印花 干练优雅

品牌盛典满减送大礼  
更多优惠券领取

¥219.00

¥ 3246

装进口 美乐宝 Apt  
amil新版爱他美奶粉 2+(2-6岁)  
00055772

DIESEL/迪赛 男式皮带手表 Q

¥98

0\*\*\*n 感觉很清  
起来很舒适

tjxiaxia 实惠，  
正品！

GAP

韩

韩

四  
办

¥98

天猫 TMALL.COM

# 我们的工作 - On Mobile



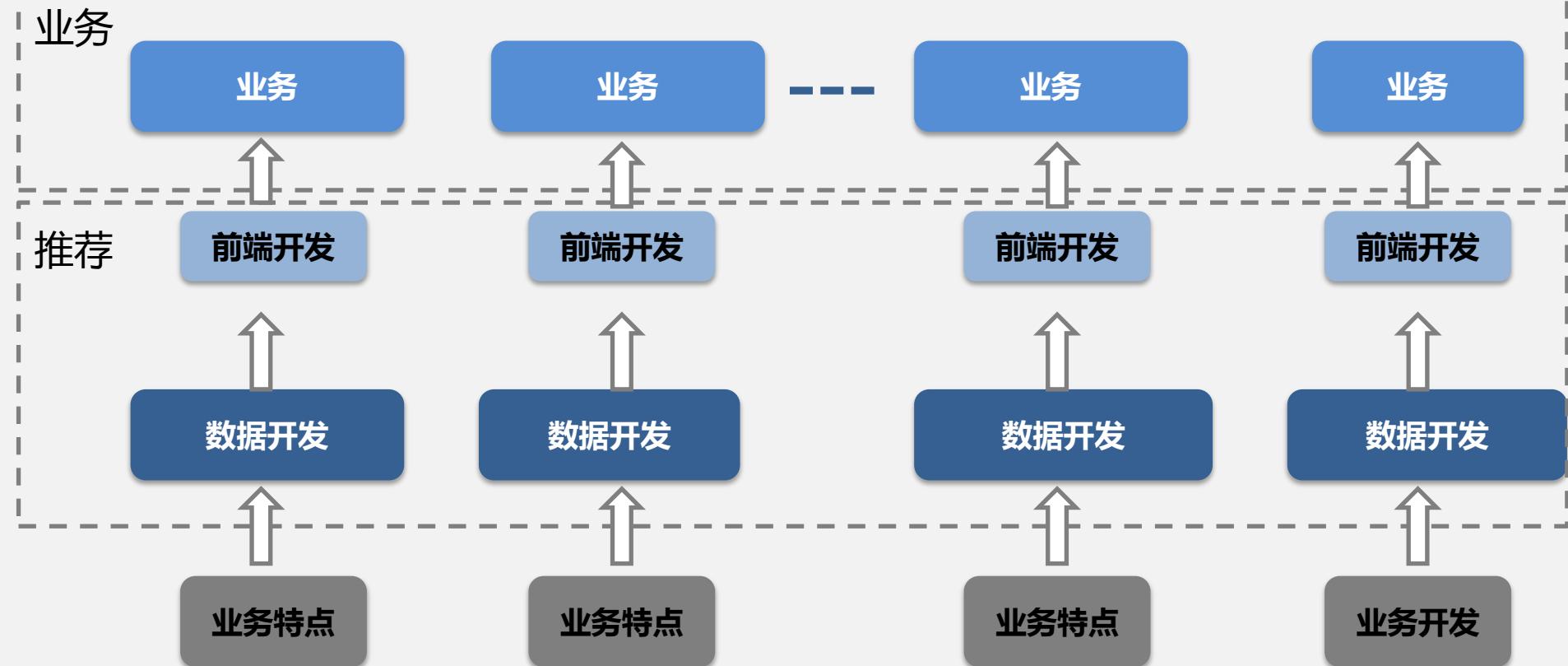
超过**60+** 的推荐实例，每日点击用户近**1千万UV**



# 我们面对的业务







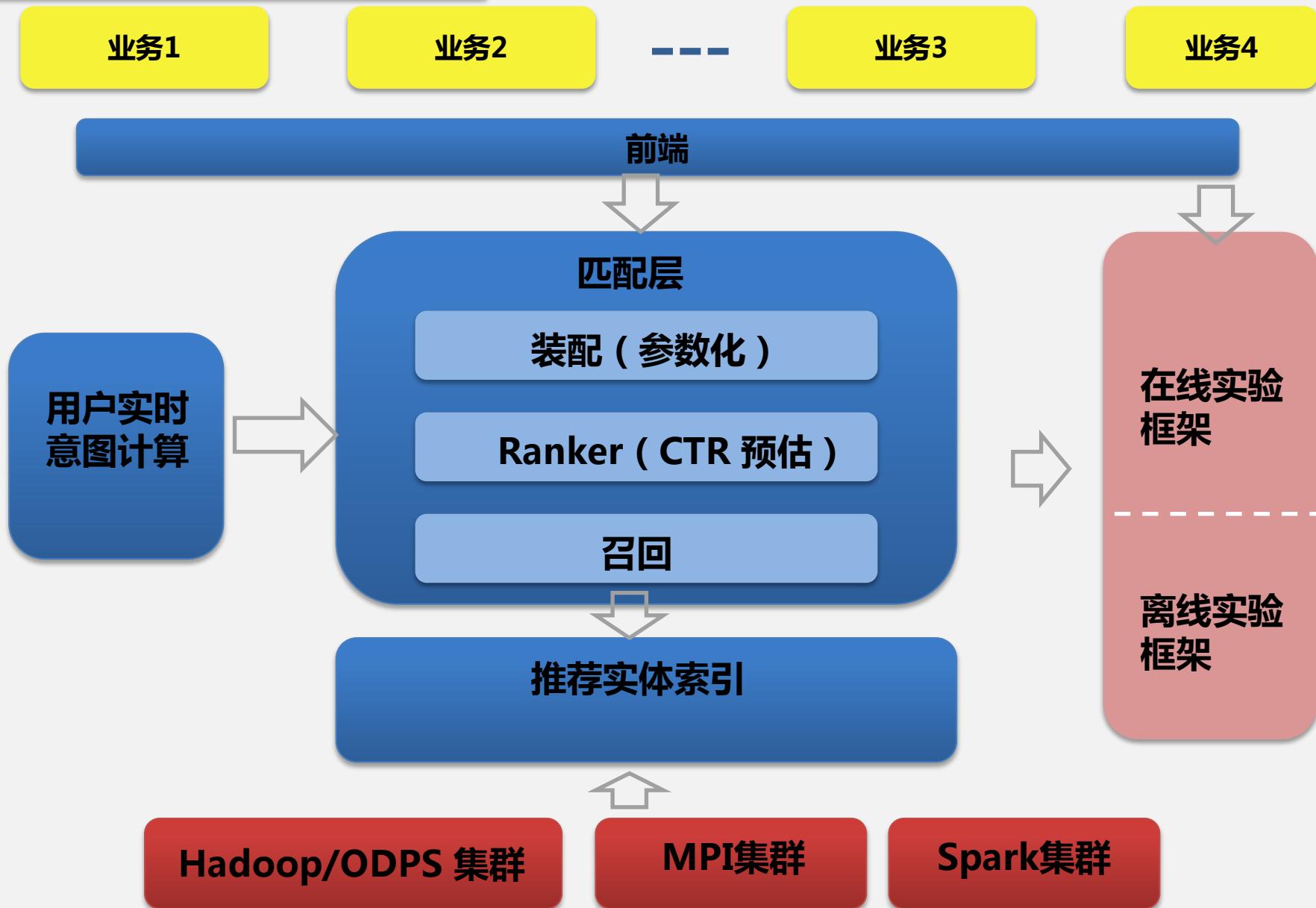
数据层笨重、逻辑层很薄、很难做到实时  
无法持续优化

数据层：数据层与业务脱离，尽量通用

逻辑层：召回与排序分离

用户意图：独立的用户意图挖掘模块

# Tmall RecSystem 1.5



推荐业务持续增长，成交占比提升**47%**



在线存储的使用量比去年同期下降**32%**



开发时间成本是原来的**1/5**

1. Online Learning

2. 集成图像技术

3. 选品



天猫推荐业务介绍/推荐系统架构历程

双11个性化会场案例

阿里巴巴大数据竞赛/天猫推荐算法大赛



2013-11-11 00:01:00

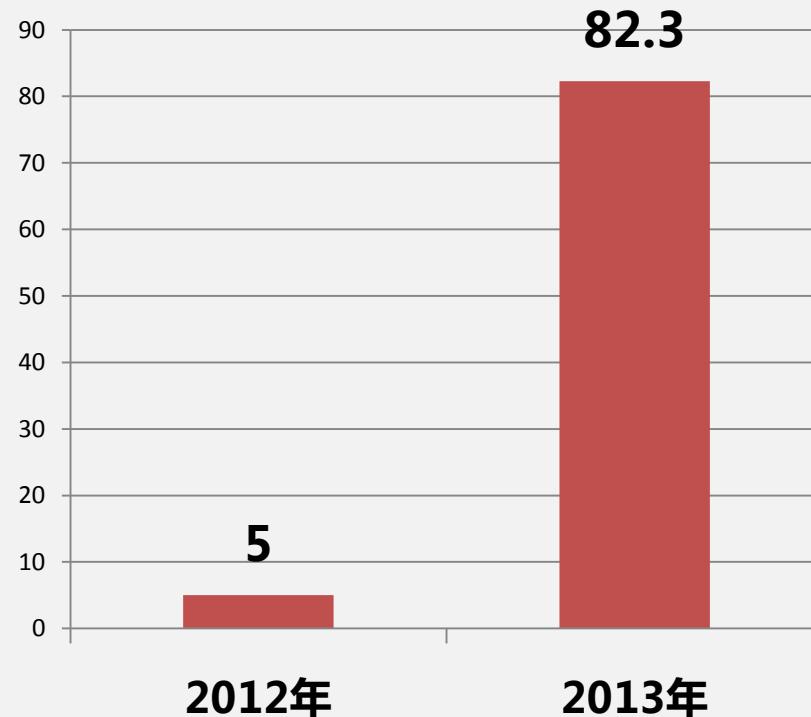


# 推荐产品双11总体贡献-成交

双11当天点击推荐产品的用户数 : **36m**

双11当天推荐引导的成交金额 : **82.3亿**

## 最近两年推荐引导成交的金额趋势图 (单位:亿)



双11当天，推荐引导的成交  
约占天猫总体的**30%**



# 具体案例



无线双11个性化会场



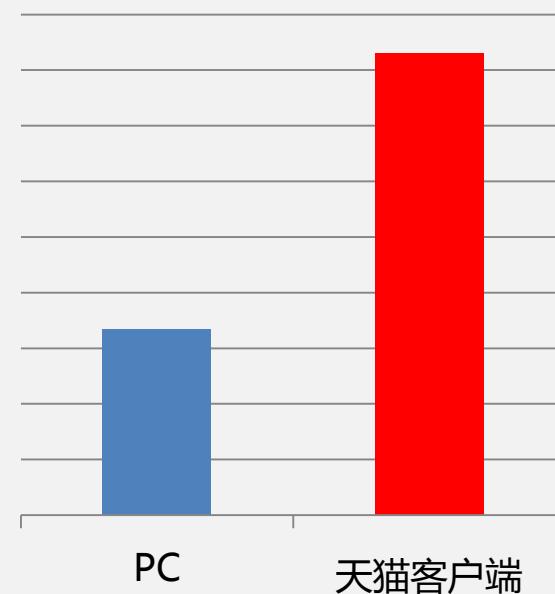
天猫 TMALL.COM

# 具体案例：

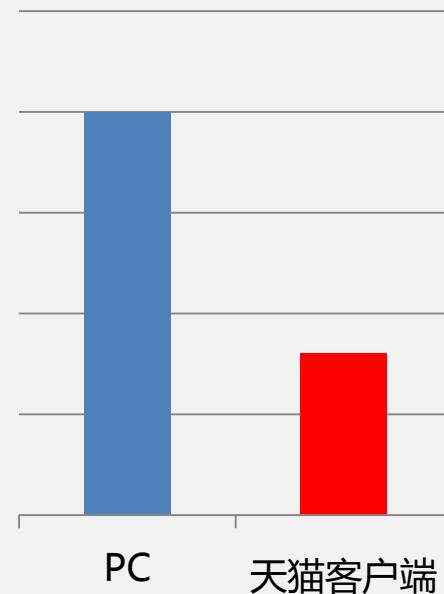


## 无线用户的特点

无线用户逛的更多



UV转化率较小



# 双11-无线个性化会场

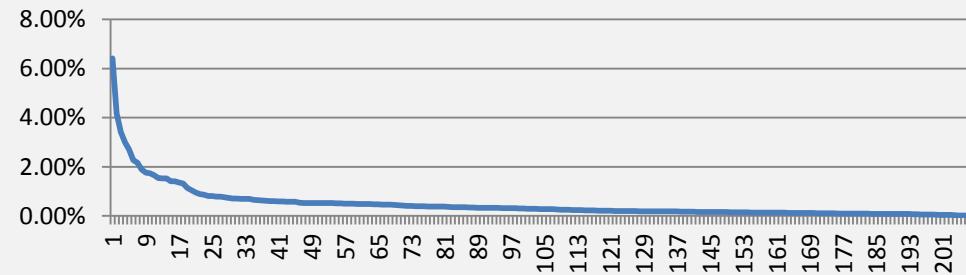


## 为什么要个性化：

女装分会场200多个品牌，在手机上要展示近100排。

点击衰减情况：

## 流量衰减情况

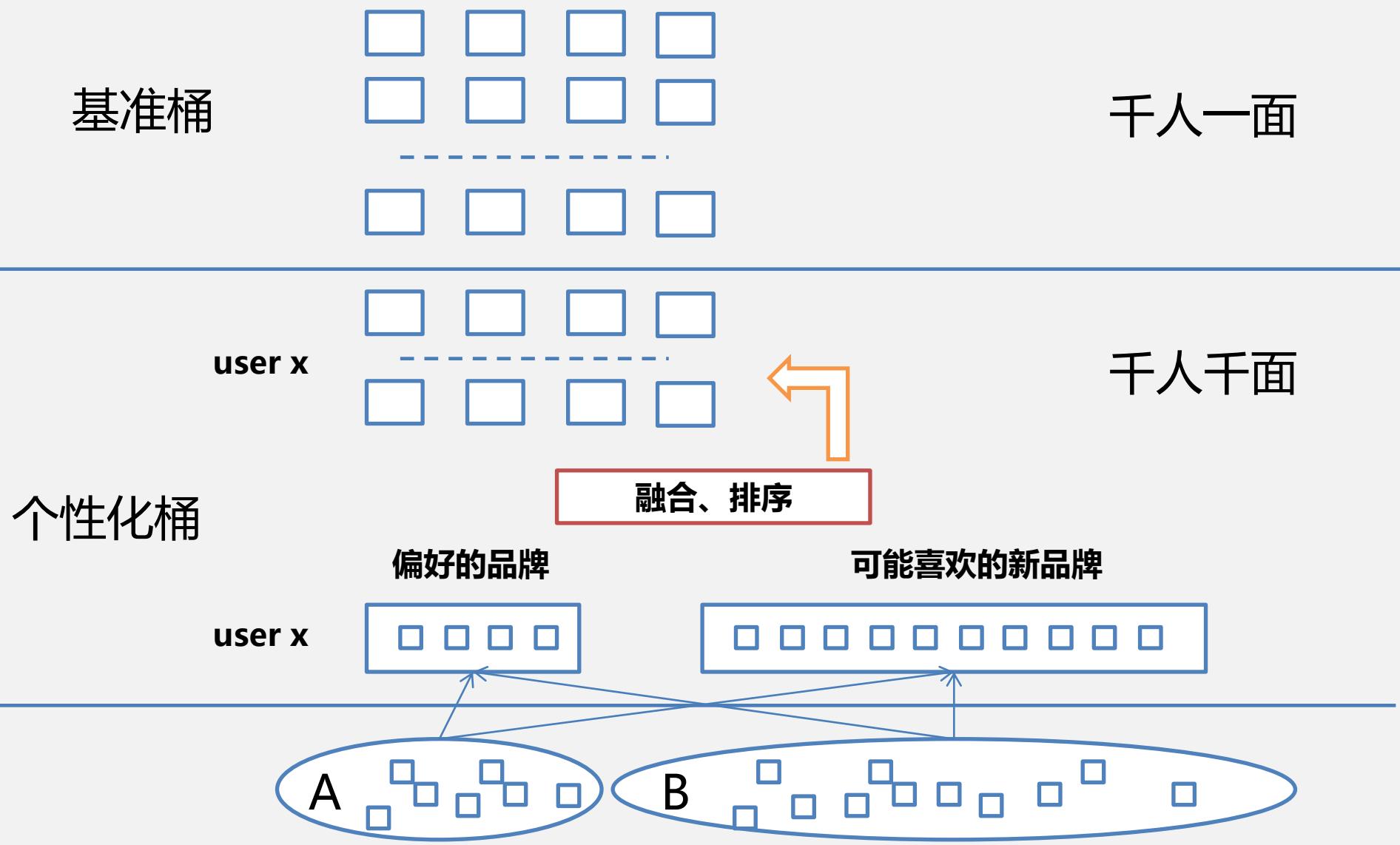


# 双11-无线个性化会场



$prob\_buy(brand_j | user_i, context)$

# 会场个性化算法-综述



我们尝试把它转化为一个点击率预估问题

Training  
Stage

	$f_1$	$f_2$	...	$f_n$	if-click
...			...		0
User <sub>i</sub> -Brand <sub>j</sub>	1	0	...	2	1
...			...		...

Prediction  
Stage

	$f_1$	$f_2$	...	$f_n$	Click prob
User <sub>i</sub> -Brand <sub>x</sub>	1	1	...	0	?

# Feature Engineering

## Feature Space :

	$f_1$	$f_2$	...	$f_n$	Click prob
User <sub>i</sub> -Brand <sub>x</sub>	1	1	...	0	?

## 类型维度：

购买

点击

## 收藏

加入购物车

### 时间维度：

1天

7天

90天

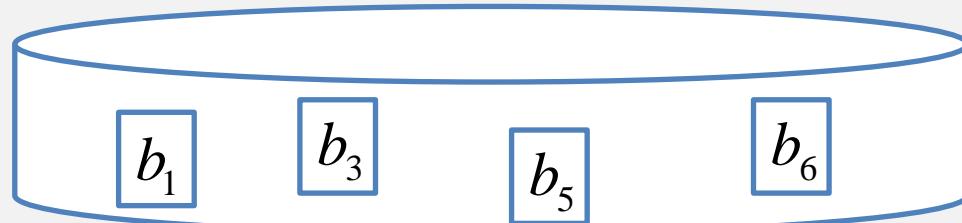
行为序列：

$b_1$   $b_2$   $b_3$



# Explore the Unknown

偏好品牌集合  
for user x



相似品牌

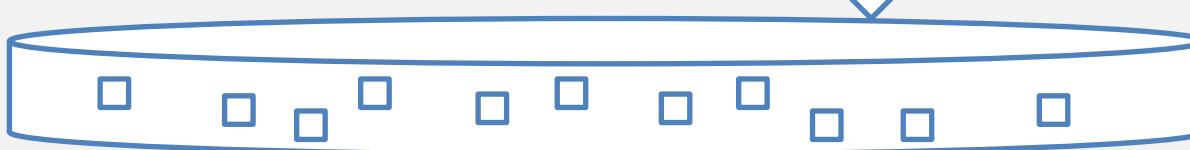
ItemBaseCF算法：

	$b_i$
$b_i$	sim



主品牌	相似品牌
$b_1$	$b_{11} ; b_{12} ; b_{13} ; \dots ; b_{1h}$
$b_3$	$b_{21} ; b_{22} ; b_{23} ; \dots ; b_{2k}$
$b_5$	.....
$b_6$	$b_{61} ; b_{62} ; b_{63} ; \dots ; b_{6k}$

新品牌  
for user x



# Models: Learning To Rank

Point-Wise

LR

Random Forest

GBDT

Pair-Wise

Rank-SVM



**女装会场**，成交金额提升**27%**

**男装会场**，成交金额提升**15%**



是一个促销的节日：**9.4亿, 52亿, 132亿, 250亿**

是一个工程的奇迹：Detail 访问次数**16亿次/当天**, 峰值吞吐率：**6.9万次/秒**

今年，开放的算法嘉年华：**开放双11的产品与算法**



天猫推荐业务介绍/推荐系统架构历程

双11个性化会场案例

阿里巴巴大数据竞赛/天猫推荐算法大赛



## 天猫推荐算法大挑战

大赛名	举办时间	队伍数	奖金数	数据
Netflix 推荐大赛	2006-2009	2000	1百万美金	48万 User ; 1.7万 Movie ; 1亿 Rating
KDD CUP 2012 腾讯资助	2012	658	8000美金	2百万微博用户; 6千 Items; 3亿曝光
Baidu 电影推荐大赛	2013	100多支	1万人民币	1.4万电影评分数据
品友互动	2013	300多支	100万人民币	

天猫推荐算法竞赛 : **7200**支队伍, **10,000**多同学, **4000**多支提交结果



## 品牌是联接商家与消费者的纽带

根据消费者的行为，计算他们对天猫品牌的偏好程度



赛题抽象

开放数据	天猫用户在2011年4月-8月的品牌行为数据： 对品牌的点击、购买、收藏、加入购物车等。
预测数据	同样这些用户在2011年9月购买的品牌



## 天猫推荐算法大挑战



内部赛：**120**多支队伍，超过**300**多名内部工程师参与

## 赛季一（淘汰赛）

时间区间：3月18日 - 4月20日

数据规模：可下载数据，小规模，包含800多个用户、10多万条记录

奖项：无奖项，从近7000支队伍中，海选最优的500支进入下一轮

## 赛季二（大数据离线赛）

时间区间：4月25日 - 7月30日

数据规模：1千多万的用户、5.7亿条记录

评奖时间：8月20日（一等奖 20万，二等奖5万，三等奖2万）

## 赛季三（共创双11）

时间区间：9月01日 - 11月11日

大奖 100万人民币



# 第一赛季参赛情况

区域	4.2报名数
中国大陆	5645
港澳台	161
海外	美国 48
	英国 7
	加拿大 6
	德国 5
	新加坡 4
	法国 4
	日本 3
	比利时 2
	荷兰 2
	芬兰 2
	爱尔兰 1
	澳大利亚 1
	韩国 1
	瑞士 1
小计 87	

截止到4月20日：7280支



清华大学  
香港科技大学  
中国科学技术大学  
中科院大学  
北邮、西交等学校做了十多次算法沙龙



今天，TOP 500 队伍正式开始第二赛季

访问 5.7亿 用户购物行为记录 @飞天

## Q&A

微博 : 2014阿里大数据竞赛技术交流

邮箱 : john.zhangq@tmall.com



# 特别感谢合作伙伴



# 特别感谢媒体伙伴 (部分)

網易 科技

腾讯视频  
V.QQ.COM

新浪微盘  
vdisk.weibo.com

懒汉互联  
汇聚/分享 [www.lanzhan.cn](http://www.lanzhan.cn)

ChinaZ.com  
China Webmaster 站长之家

ZDNet.com.cn  
云计算第一门户

**51CTO.com**

网易云阅读

华章科技  
HZ Books

w3tech

OSGi  
中文社区

TURING  
图灵教育

福布斯中文网  
**Forbes**  
FORBESCHINA.COM

腾讯精品课

第1财经  
C B N

ifeve

FT  
中文网  
[ftchinese.com](http://ftchinese.com)

动点科技  
[cn.technode.com](http://cn.technode.com)

腾讯大讲堂  
DJT.QQ.COM

手机腾讯网  
4G.QQ.COM

TECH  
开发者社区

DOIT  
中国 IT 新媒体

通信产业报  
COMMUNICATIONS WEEKLY

IT168.com  
[www.it168.com](http://www.it168.com)