**专 利 申 请 技 术 交 底 书**

**专利卷号（法务部填写）：**

**如有疑问，请咨询集团法务部。**

**1688事业部/国际事业部/商业智能部1688DW/ 技术质量部1688技术质量员工请联系李海波，分机：39508，Email:** [ricky.lihb@alibaba-inc.com](mailto:ricky.lihb@alibaba-inc.com)**；**

**淘宝/天猫事业部/聚划算事业部/航旅事业部/淘点点事业部/技术质量部共享业务员工请联系季江胜，分机：39694，Email:** [jiangsheng.jijs@alibaba-inc.com](mailto:jiangsheng.jijs@alibaba-inc.com)**；**

**一淘及搜索事业部/无线事业部/来往事业部/旺旺与客户端事业部/阿里妈妈事业部/技术质量部阿里妈妈技术/数字娱乐事业部员工请联系宋家扬，分机：39832,Email:** [jiayang.songjy@alibaba-inc.com](mailto:jiayang.songjy@alibaba-inc.com)**;**

**小微金服员工请联系潘葚（周泰），分机：37079，Email:** [shen.pans@alibaba-inc.com](mailto:shen.pans@alibaba-inc.com)**;**

**阿里云事业部/云OS事业群/集团技术保障/数据平台事业部/共享业务事业部/菜鸟/技术质量部中间件及大数据员工请联系岳强，分机：39750，Email:** [qiang.yueq@alibaba-inc.com](mailto:qiang.yueq@alibaba-inc.com)**。**

**以下内容请发明人填写：**

**发明名称：\_\_\_\_一种基于用户关系的评论推荐的方法和装置\_\_\_\_**

**发明人姓名：\_\_\_\_\_\_\_李生、谢朋峻、孙常龙、郎君\_\_\_\_**

**发明人花名：\_\_\_\_\_\_\_子尹、程辰、常龙、上乘**

# 发明人工号：­ \_\_\_\_083267、051276、041441、071372\_\_\_

**第一发明人身份证号码（**必填项，针对中国发明人**）：\_320925198909146456\_\_\_\_\_\_\_\_\_\_**

**技术联系人：\_\_李生\_\_\_\_**

**电话：\_\_\_\_15501682387\_\_\_**

**E-MAIL：\_\_lisheng.ls@alibaba-inc.com\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**所属公司/部门\_\_\_搜索事业部-Data Alibaba-语言技术\_\_\_**

**注：1、以上内容均需要填写。**

**2、“发明人花名”一栏中，淘宝发明人必填。**

**术语解释：**

1. **本发明是否已在或计划在产品中实施？如果是，请列出产品名。是否已经或计划对外书面公开？如果是，请具体说明。**

计划实施。

名称:基于用户关系的评论推荐。即将在商品评论页呈现。

1. **本发明是否有敏感的部分不适合作为专利申请公开？例如，会不会对黑客或竞争对手提供优势？**

没有敏感地方。

1. **本发明要解决的技术问题是什么？**

依据用户之间的关系（如淘友，同城，身材相似等），解决商品评论的可信度问题。

**4、详细介绍技术背景,并描述已有的与本发明最相近似的实现方案**（包括两部分：1、作为本发明基础的且帮助理解本发明公知技术内容；2、与本发明最接近的已有技术方案的说明（对于方法，应说明现有方法的步骤，对于装置，应当说明结构组成及其关系））

# 商品评论信息可以给用户提供购前决策信息，维持一个良好的评论环境对买家和卖家都是至关重要的。但目前评论信息依然缺乏可信性，主要是刷评，广告信息占领了评论。

# 目前评论排序从一定程度上治理了糟糕的评论环境，这方面工作包括评论折叠和评论垃圾识别等。

# （a）评论折叠：评论中有很多刷单或恶意用户进行重复评论，折叠工作让重复的评论只展现一次。

# （b）评论垃圾识别：主要针对垃圾评论（比如广告、涉黄、涉政等）进行过滤。

# 但是由于评论用户匿名，呈现给用户信息较少，评论的价值还有待商榷。

# 我们尝试利用社交关系、用户个性化信息等，计算评论用户和当前浏览用户的相关性并给评论用户打上一定的标签（身材类似，同城，年龄相仿，趣味相同等），这样可以从另一维度上解决评论可信性问题。

**5、以因果关系推理的方式推导出现有技术的缺点是什么？针对这些缺点，说明本发明的目的及能够达到的技术效果。**（现有技术的缺点是针对于本发明的优点来说的，本发明不能解决的缺点不必写）

商品评论信息可以给用户提供购前决策信息，维持一个良好的评论环境对买家和卖家都是至关重要的。目前商品评论排序工作已经可以解决部分问题，比如，评论折叠和评论垃圾识别等。

**评论排序工作从一定程度上整治了评论的环境，但是商品评论依然存在不可信问题。由于评论褒贬不一，给用户造成困扰。**

**无论是评论折叠和垃圾识别都是从评论文本这单一维度进行考虑，这方面的工作并不能解决评论信息可性度问题。具体来讲：**

**（a）评论折叠主要是针对重复评论进行治理。或是刷评或是恶意诋毁，部分用户进行虚假地发布同一评论或高度相似的评论。为了保持良好的评论氛围，折叠工作将重复评论或高度相似评论仅透出一条。**

**（b）评论垃圾识别主要是针对广告信息、涉黄、涉政等垃圾信息进行过滤。为避免虚假交易和恶意竞争以及散播反动信息，必须将这些信息过滤掉。**

**以上两个工作使得评论环境得到优化，但是并不能解决用户需求：评论的可信度。目前淘宝的评论（如图1）对评论用户进行匿名，仅仅展示评论的内容我们仍为不具有太大的可性度。**

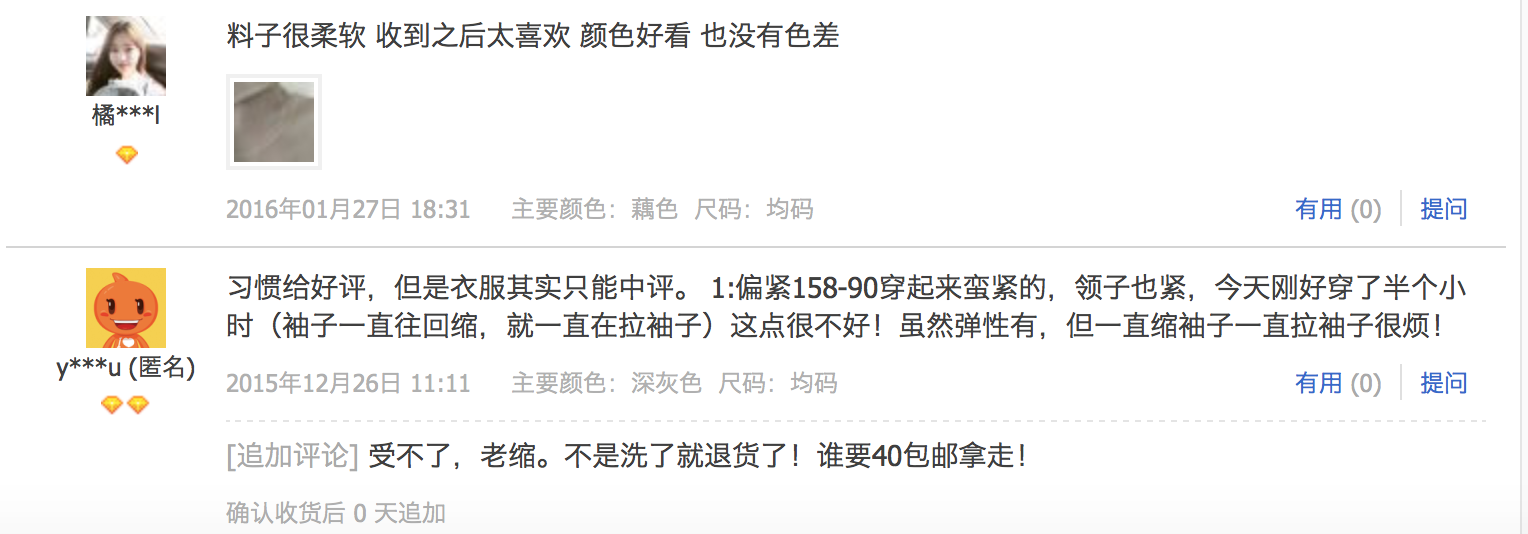
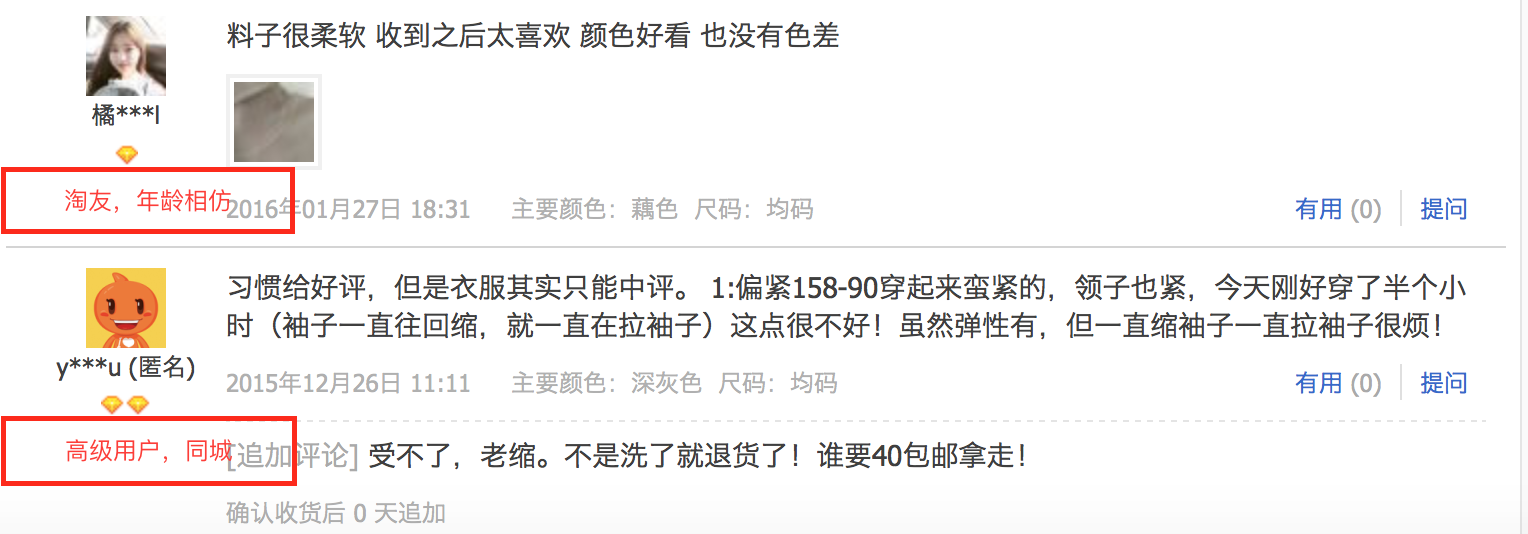


图1：淘宝现有商品评论截图

**本发明旨在通过用户与用户之间丰富的关系，给评论的用户赋予一定的标签（如身材类似，同城，达人等），同时结合用户自身信息对呈现给用户的评论进行个性化推荐。我们期望呈现给用户多方面的信息，提高评论的价值。**

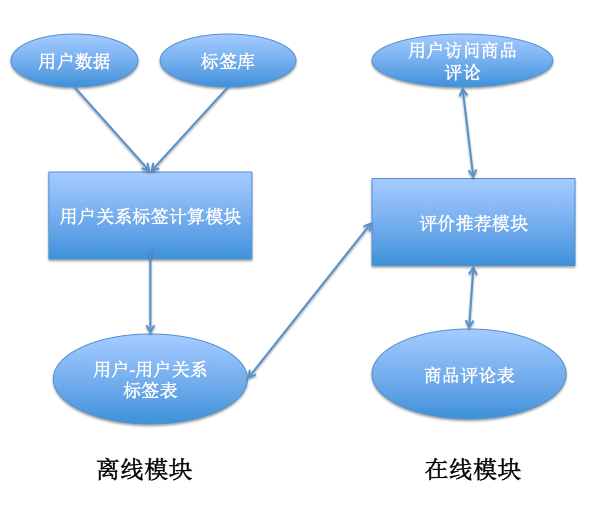
图2：基于用户关系的淘宝评论展现



**图2是基于用户关系推荐后的淘宝评论，在展现时给用户了提供了其他维度的信息，比如评论的用户和当前用户是淘友关系等，这样可以从其他维度给用户提供决策信息。**

**6、本发明技术方案的详细阐述，应该结合流程图、原理框图、电路图、系统结构图进行说明**（发明中每一功能的实现都要有相应的技术实现方案；所有附图都应该有详细的文字描述；方法专利都应该提供流程图，并提供相关的系统装置。5.1 本部分为专利申请最重要部分，需要详细提供；5.2 发明必须是一个技术方案，不能只有原理，也不能只做功能介绍；5.3 对于软件、业务方法，要提供流程图；5.4 必须结合流程图、原理框图、电路图、系统结构图等附图进行说明。）

I 系统框图

****

**图3：基于用户关系的评论推荐的系统框图**

II. 基本流程

1. **离线计算部分**

**模块1： 用户关系计算模块**

**（a）关系库：关系库主要定义了一些关系的类型，用于在线评论推荐时作为推荐理由。主要包含：“淘宝好友”、“同小区用户”、“相同购物偏好”、“身材相似用户”等；**

**（b）用户数据：用户数据包含了淘宝用户关注关系、用户的人口统计学信息、用户的购物行为数据等；**

**（c）用户关系标签计算模块：该模块主要针对标签库不同的关系类型，基于用户数据来计算用户之间的关系。**

**以“身材相似用户”关系为例，我们基于用户数据能够获取不同用户的身高、体重，基于此，我们将身高差距小于等于2厘米，而且体重差距小于5KG的用户之间标记上“身材相似用户”；类似的，我们基于用户住址信息可以标记出“同小区用户”；基于用户的购物历史，可以挖掘出“相同购物偏好用户”；基于淘宝好友关注信息，可以标记处“淘宝好友”。最终得到的表如表1所示：**

|  |  |  |
| --- | --- | --- |
| 关系类型 | 用户ID | 用户ID |
| 同小区用户 | uid1 | Uid4 |
| 淘宝好友 | Uid2 | Uid4 |
| 相同购物偏好 | Uid11 | Uid12 |

表1：用户-用户关系表

1. **在线推荐部分**

**当用户Uid1访问某个商品评论时，评价推荐模块首先去访问商品的评论数据，获取默认的评价表，如表2所示：**

|  |  |
| --- | --- |
| 用户ID | 评论内容 |
| Uid2 | 羽绒服很暖和，快递很快 |
| Uid3 | 羽毛会掉出来，质量不是很好 |
| Uid4 | 款式不错，穿着很暖和 |

**表2：商品评论原始表**

**同时访问表1的用户-用户关系表，我们得到Uid1和Uid4属于“同小区用户”关系，于是将Uid4对应的评论“款式不错，穿着很暖和”排序提前，同时打上标签“同小区用户”，最终线上推荐结果如表3所示。**

|  |  |  |
| --- | --- | --- |
| 用户ID | 关系标签 | 评价内容 |
| Uid4 | 同小区用户 | 款式不错，穿着很暖和 |
| Uid2 |  | 羽绒服很暖和，快递很快 |
| Uid3 |  | 羽毛会掉出来，质量不是很好 |

**表3：关系上提后的评价内容**

**真实线上效果大概如下图所示：**

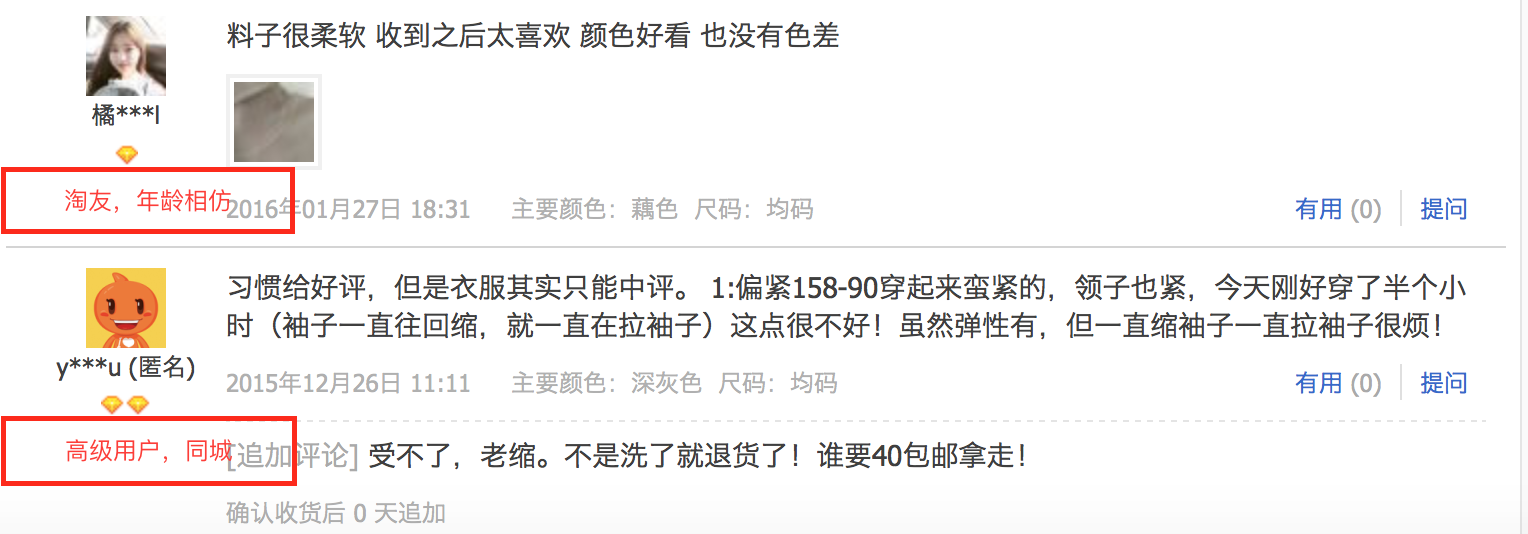


图4：基于用户关系的评论推荐的效果

**7、在实施本发明中，是否有某种计算机语言，软件或硬件能够达到最佳的实施效果？**

**离线数据的统计牵涉到大规模的数据处理，所以需要大规模并行计算的需要，需要云计算平台。同时需要支持实时查询的keyvalue系统。**

**8、本发明的关键点和欲保护点是什么？**（请提炼出本发明的技术创新点，以提醒专利代理人注意，便于其撰写权利要求书）*。*

依据用户之间的丰富关系（如淘友，同城，身材相似等），解决商品评论可信度问题。

**9、能实现本发明目的的其他替代方案**（尽量多列举可实现本发明的的其他替代方案，以便充分公开本发明）

注意：感谢您对阿里巴巴集团专利工作的支持和所付出的辛勤工作，我们将按照以下条件对您的提案进行审核，通过审核后我们将按照您所属BU的《专利奖励办法》中的规定，发放相应的奖励。

审核条件：

-是基础技术还是周边技术；

-是否是公司重点项目所涉及的技术；

-是否解决了业界重大的技术问题；

-是否业界有近似产品或技术；

-是否代表业界未来发展方向；

-是否易于发现竞争对手的侵权行为及证据；

-是否满足专利法的相关规定。