

# 内容推荐系统和方法

申请号：[201110339401.6](#)

申请日：2011-10-24

申请(专利权)人	<a href="#">微软公司</a>
地址	美国华盛顿州
发明(设计)人	<a href="#">D 德希尔</a>
主分类号	<a href="#">G06Q30/02(2012.01)I</a>
分类号	<a href="#">G06Q30/02(2012.01)I</a>
公开(公告)号	102446330A
公开(公告)日	2012-05-09
专利代理机构	<a href="#">上海专利商标事务所有限公司</a> 31100
代理人	<a href="#">陈斌</a>



## (12) 发明专利申请

(10) 申请公布号 CN 102446330 A

(43) 申请公布日 2012. 05. 09

(21) 申请号 201110339401. 6

(22) 申请日 2011. 10. 24

(30) 优先权数据

12/911, 671 2010. 10. 25 US

(71) 申请人 微软公司

地址 美国华盛顿州

(72) 发明人 D·德 希尔

(74) 专利代理机构 上海专利商标事务所有限公司  
31100

代理人 陈斌

(51) Int. Cl.

G06Q 30/02 (2012. 01)

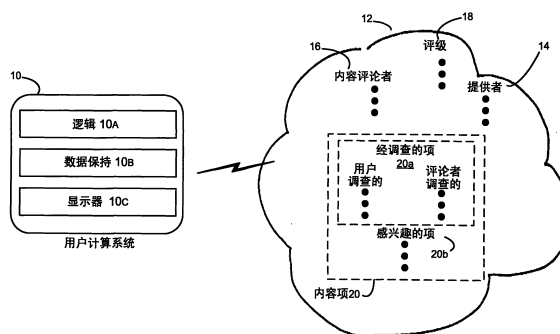
权利要求书 2 页 说明书 5 页 附图 3 页

### (54) 发明名称

内容推荐系统和方法

### (57) 摘要

本发明涉及内容推荐系统和方法。一种推荐并评级内容和 / 或生成内容消费者和评论者之间的确定的密切关系的系统和方法。所述方法包括确定一个或多个用户调查的内容项的评级, 并为多个内容评论者的每个确定一个或多个评论者调查的内容项的评级。随后, 在共同调查的内容项的用户评级和评论者评级之间执行比较。结果, 建立用户和内容评论者之间的密切关系, 并且为用户提供由高密切关系内容评论者产生的评论和评级。



1. 一种预测用户与内容评论者 (16) 的偏好的密切关系的方法 (200), 所述内容评论者的评论 (18) 可通过计算机化通信网络 (12) 访问, 所述方法包括:

为用户确定 (202) 一个或多个用户调查的内容项 (20a) 的评级 (18);

对于多个内容评论者 (16) 中的每个, 确定 (204) 一个或多个评论者调查的内容项 (20a) 的评级;

对于一个或多个共同调查的内容项 (20a), 将所述一个或多个用户调查的内容项 (20a) 的评级 (16) 与所述一个或多个评论者调查的内容项 (20a) 的评级 (18) 进行比较;

基于所述比较步骤, 确定 (210) 所述用户与所述多个内容评论者 (16) 中的每个的密切关系;

对于正被所述用户考虑的或可以作为给所述用户的潜在推荐的感兴趣的选中项 (20), 为所述用户提供 (212) 来自所述多个内容评论者 (16) 中的所选数目的内容评论者的评论 (18), 所选数目的内容评论者 (16) 被确定为相对高的或最高密切关系。

2. 如权利要求 1 所述的方法, 其特征在于, 确定所述用户与所述多个内容评论者的每个的密切关系包括在所述一个或多个共同调查的内容项上执行均方操作。

3. 如权利要求 1 所述的方法, 其特征在于, 对于所述多个内容评论者的每个, 所述比较包括对所述一个或多个共同调查的内容项的归一化评级的比较。

4. 如权利要求 3 所述的方法, 其特征在于, 对于所述内容评论者中的每个, 所述比较包括对所述一个或多个共同调查的内容项上执行均方操作。

5. 如权利要求 1 所述的方法, 其特征在于, 进一步包括向所述用户提供正被所述用户考虑的或可以作为给所述用户的潜在推荐的感兴趣的所述选中内容项的合成评级, 其中使用已经评论了正被考虑的所述选中项的所述内容评论者中的一个或多个的经密切关系加权的评级来生成所述合成评级。

6. 如权利要求 1 所述的方法, 其特征在于, 所述为所述用户确定一个或多个用户调查的内容项的评级包括在缺少明确用户评级的情况下推断所述一个或多个用户调查的内容项的评级。

7. 如权利要求 6 所述的方法, 其特征在于, 所述推断是响应于对所述一个或多个用户调查的内容项的用户发起的展示来执行的。

8. 如权利要求 1 所述的方法, 其特征在于, 正被所述用户考虑的所述选中的内容项是数字媒体项。

9. 一种被配置成预测用户与内容评论者 (16) 的偏好的密切关系的计算系统 (10), 所述内容评论者的评论 (18) 可通过计算机化通信网络 (12) 访问, 所述计算系统包括:

数据保持子系统 (10b), 包含可由处理器 (10a) 执行的指令, 所述指令被执行以:

为所述计算系统 (10) 的用户确定一个或多个用户调查的内容项 (20a) 的评级 (18);

对于多个内容评论者 (16) 中的每个, 通过与所述计算机化通信网络 (12) 的交互来获得一个或多个评论者调查的内容项 (20a) 的评级 (18);

对于一个或多个共同调查的内容项 (20a), 将所述一个或多个用户调查的内容项 (20a) 的评级 (16) 与这样的内容评论者 (16) 的该一个或多个评论者调查的内容项 (20a) 的评级 (18) 进行比较;

基于所述比较步骤, 确定所述用户与所述多个内容评论者 (16) 的每个的密切关系, 其

中这样的密切关系确定是使用对所述一个或多个共同调查的内容项 (20a) 的均方根操作来执行的 ; 以及

对于正被所述用户考虑的或将推荐给所述用户的选中的内容项 (20) 并基于所述用户和所述多个内容评论者 (16) 之间的所述密切关系来从所述多个内容评论者 (16) 中的所选择的一些内容评论者中获得所选中的内容项 (20) 的评论, 其中所述选择是基于所述用户和所述内容评论者 (16) 之间的密切关系的。

10. 如权利要求 9 所述的方法, 其特征在于, 进一步包括向所述用户提供正被所述用户考虑的或将推荐给所述用户的所述选中内容项的合成评级, 其中所述合成评级是使用已经评论了正被考虑的所述选中内容项的内容评论者中的一个或多个的经密切关系加权的评级来生成的。

## 内容推荐系统和方法

### 技术领域

[0001] 本发明涉及内容推荐系统和方法。特别是,提供一种推荐并评级内容和 / 或生成内容消费者和评论者之间的确定的密切关系的系统和方法。

### 背景技术

[0002] 因特网和其它通信网络在它们提供将用户与相关内容进行连接的解决方案方面变得日益完善。尽管这样,因特网和其它网络还是构成了大量杂乱无序的信息,对于用户来说要从中标识出确保是有用的和 / 或感兴趣的信息还是非常困难的。存在各种解决这个问题的方法,包括协作过滤和各种社交推荐工具。

[0003] 例如,许多电子商务网站使用基于用户在站点上展示的特定物品之间的浏览将备选的或补充的产品推荐给用户的功能。在许多情况下,这些推荐是基于由网站累积的购买者数据,例如示出购买了第一产品的用户还可能想要购买第二相关产品的数据。这样的方法的一个问题在于不存在一种建立给定用户将具有与网站的其他用户相类似的偏好的机制。因此,所述推荐可以是高度推测性的,并且因而,这些系统提供相关推荐给用户的能力受到了限制。

### 发明内容

[0004] 本发明提供了一种推荐内容和 / 或生成内容消费者和评论者之间的确定的密切关系的系统和方法。在一个方面,本发明涉及一种用于预测用户与内容评论者的偏好的密切关系的方法。所述方法包括为用户确定一个或多个用户调查的内容项的评级。随后,对于多个内容评论者中的每个,确定一个或多个评论者调查的内容项的评级。随后,将用户调查的和评论者调查的内容项的评级进行比较,通常是对这两个调查实体之间重叠的那些内容项的评级进行比较。作为比较的结果,建立用户和内容评论者之间的密切关系。随后,对于用户正在考虑的选中项或对于用户可能感兴趣的项,为用户提供可以包括评级的由高密切关系的内容评论者所产生的推荐和 / 或评论。而且,提供执行各种步骤的计算系统,所述计算系统可以包括为了标识高密切关系评论者并提供高质量推荐和评级或其他评论给所述用户的上述那些步骤和 / 或其他步骤。

[0005] 提供本发明内容以便以简化形式介绍在以下具体实施方式中进一步描述的一些概念。本发明内容并不旨在标识所要求保护主题的关键特征或必要特征,也不旨在用于限制所要求保护主题的范围。此外,所要求保护的主体不限于解决在本公开的任何部分中提及的任何或所有缺点的实现。

### 附图说明

[0006] 图 1 是根据在此所述的示例的可在其中结合内容推荐系统和方法的示例计算系统和环境。

[0007] 图 2 是示出密切关系的示例。

[0008] 图 3 描述示出经调查的内容项的示例。

[0009] 图 4 是用于生成密切关系评估并通过提供评级、其他评论和内容推荐来响应所述评估的示例方法。

## 具体实施方式

[0010] 本发明涉及用于预测用户对电子地可访问内容项（例如视频和其他数字媒体）的兴趣并标识具有与特定用户相类似的偏好的内容评论者的系统和方法。在许多情况中，这些内容评论者的标识被用于改善提供给用户的推荐的质量。在此的许多示例中，将把用户描述为具有与例如职业电影评论家、电子商务站点的同伴购买者等在线内容评论者的密切关系。当在此使用时，“在线”仅仅意味着所标识的信息和实体是彼此通信并且电子地可访问的。因此，所述术语可以包含例如因特网等公共网络或网络集合以及私有或受管网络。实际上，“在线”关注能够让用户接收内容推荐和消费电子内容项的任意类型的系统或系统的集合。

[0011] 图 1 描述了在其中可以使用在此描述的系统和方法的示例环境。用户可以使用例如示例性的计算系统 10 等计算设备来搜索并获得对可通过计算机化的通信网络 12 获得的内容项的访问，所述网络 12 可以是因特网或启用对内容项访问的另一系统 / 网络。

[0012] 如图所示，网络 12 可以包括或启用对提供内容的多个内容提供者 14、内容评论者 16 以及可用内容项的评论 / 评级 18 的访问。如在此所述，术语“评论”可以指或包括能够让用户决定他们是否对获得或消费内容项感兴趣的任意类型的信息。例如，评论可以包括内容项的书面描述（例如包含观点、评语等）或它可以包括项的评级（例如从一颗到五颗星或其他的数字或其他评估）。内容提供者 14 可以是内容项 20 的源或位置，所述内容项 20 可以包括经调查的项 20a（如下所述）和用户感兴趣的项 20b。在该所描述的示例中所关注的内容项可以是音乐或视频文件（可下载的或流传送的）、其他类型的数字媒体和 / 或可以使用计算机来电子地访问的任意其他类型的内容。

[0013] 在第一示例中，所描述的系统可以用于预测用户对正在考虑的内容项，例如内容项 20b 之一，的兴趣。例如，用户可能正在考虑是否购买可流传输的电视剧集、音乐文件、播客（podcast）等。在该第一示例中，通过访问用户对一个或多个内容评论者 16 的密切关系来预测兴趣。具体而言，如果用户对某个内容项给予很高评级，可以认为他们赞成已经给予这些相同的项很高评级的在线内容评论者。

[0014] 为了评估密切关系，在此的方法关注对内容项的调查。特别地，接收或推断对已经展示给用户的一个或多个内容项的用户评级。用户可以已经购买了这些项、已经浏览表征这些项的网页、已经购买或评论了紧密相关的项等。这些项目可被称为用户调查的项。所述经调查的项构成了使得能够知道用户的偏好的累积信息。

[0015] 如上简明示出的，评级可以是明确的和 / 或推断的。明确评级可以包括熟悉的“星评级”系统，在其中，用户根据他们对特定内容项的观点来分配相对较多或较少数目的星。例如，用户可以给他不喜欢的内容项分配一颗星。在另一示例中，用户可以给他喜欢的内容项分配五颗星。推断评级是从较不明确的行为中生成的计算得到的评级。例如，用户可以展示由用户发起的展示行为，例如选择并查阅表征该项的网页、选择购买 / 消费紧密相关的项、所选择的查看和购买的数目和 / 或频率，等等。推断的强或弱取决于评论者与该

类型的内容的关系,或基于该内容项与其他内容项的相关度。推断 / 计算得到的评级还可以根据特定行为变化。例如,当用户仅仅浏览表征该项的页面时可以分配四颗星中的三颗星。另一方面,当购买项时,推断可以生成更高的评级,例如四颗星。在另一个示例中,如果用户要购买某种构造 / 型号的照相机,可以认为相似的评级会附加给来自同一制造商的其他类似的照相机。在又一个示例中,如果用户对具有特定演员的电影进行明确评级,则类似的评级可以通过推断分配给具有相同演员的其他电影。在用户音乐库中,可以将更高的评级分配给已经被更多次数聆听的歌曲。又一示例是当用户观看整个视频剪辑时可以推断高的评级 / 兴趣。如上述示例所表明的,可以使用各种方法来生成用户调查的项的推断评级。

[0016] 还可以针对多个在线内容评论者,例如评论者 16,累积调查评级。当这些评论者产生对由用户调查的项的评级时,可以进行评级比较以得到用户 - 评论者密切关系评级。典型地,内容评论者的评级是明确的,但也可使用推断评级。例如,如果职业电影评论家一致对来自特定导演的电影评级很高,那么即使评论家没有明确地评论来自该特定导演的另一电影,也可为该另一电影附加类似的高的评级。而且,与用户一样,推断的强或弱取决于评论者与该内容项的关系,或基于该内容项与其他内容项的相关度。例如,可以假设与演员表成员相比较,电影的导演对于电影质量来说具有更强大的决定性。在另一方面,通过分析,可以确定特定用户实际上在特定演员存在的这个因素上赋予比其他因素更重的加权。同样,可以通过各种方法来得到推断。

[0017] 在任意情况中,一旦获得所调查的项的推断 / 计算或明确的评级,可以将对用户和给定在线内容评论者共同的所调查的项进行比较以评估用户 - 评论者密切关系。在比较和得出密切关系的决定中可以使用各种方法。例如,在大量评论者数量的情况下,可以丢弃(假设零密切关系)分配不同的评级(例如较低评级)的内容评论者。在另一方法中,就在线内容评论者对调查项的评级与该用户的评级不同而言,他们的偏好将被折扣。一种可以使用的方法是均方操作,在该操作中,对分数差值应用均方计算以对分配给特定内容评论者的加权和密切关系进行折扣。例如,如果用户对项分配了四颗星,均方方法可以将两颗星的评论评级折扣得比三颗星评论评级更加深入。还有,可以理解,在一些情况中,归一化评级以适应不同的评级方案是合乎需要的(例如,提供对以五颗星尺度来评级的项和以四颗星尺度来评级的项的精确比较)。

[0018] 无论如何,可以以各种方式来使用得到的密切关系。例如,在用户考虑特定感兴趣的项(例如项 20b)的情况下,可为用户提供仅来自的具有非常高的用户 - 评论者密切关系分数的评论者的对该项的评论。在一个非限制性示例中,可为用户只提供项的两个评论,但是这两个评论将来自与已经为该用户建立的口味偏好最紧密地符合的评论者(例如职业电影评论家)。

[0019] 在另一个示例中,使用该系统以对正被考虑购买、消费等的项 20b 的用户兴趣水平进行预测。这一预测可以涉及该项的合成评级,其中来自多个内容评论者的评级被合并成该合成评级。一个方法是使用已经在用户和已经评论 / 评级该项的评论者之间建立的密切关系来加权该评级。相对于先前的示例,可为用户简单地提供合成评级以取代从在线内容评论者接收实际的评论,虽然应该理解评级本身从某种意义上来说是一种“评论”。当然,可以组合这些方法,以使用户将接收合成评级和来自高密切关系的评论者的对正被考虑的项的描述评论。

[0020] 图 4 提供了由本发明关注的示例方法的流程图描述。在方法 200 (图 2) 中,通过首先确定由用户调查的一个或多个用户调查的内容项的评级来完成用户与内容评论者的密切关系,如在 202 所示。如在本发明的其他部分所解释的,“经调查的”意味着用户或评论者以一种使该系统能够确定用户是喜欢还是不喜欢一数字内容项并且还确定偏好的强度的方式来与该项进行了交互。例如,评级一组歌曲意味着用户已经调查了这些歌曲。完整地观看视频也可以被认为是视频的一种“调查”,因为它使得该系统能够假设用户喜欢该视频。继续图 4,在 204,为多个内容评论者的每个确定一个或多个评论者调查的内容项的评级。如前所述,在步骤 202 和 204 中的评级可以是明确的或推断的。随后,在 206 和 208,将对共同调查的内容项的用户评级和评论者评级进行比较以确定用户-评论者密切关系。当然,这些步骤可以是迭代的以考虑多个用户和多个评论者情况。最后,对于用户正考虑的选中项,可为用户提供来自高密切关系的在线内容评论者或相对高密切关系的在线内容评论者(具有比其他在线内容评论者更高的密切关系)的对正在考虑的项的选中评论,如在 210 和 212 所示。另外地或替换地,该方法还可以包括为用户提供正在考虑的项的评级,如在 212 所示。方法还可以包括向用户提供高质量推荐,并且不局限于提供已经由用户评论或浏览的项的评估。如上所述,这样的评级可以从来自多个评论者的评级来生成,和/或可以是使用用户-评论者密切关系的加权来生成的合成评级。

[0021] 在某些实施例中,以上所描述的方法和过程可应用于包括一个或多个计算机的计算系统。特别是,此处描述的方法和过程可被实现为计算机应用、计算机服务、计算机 API、计算机库和/或其他计算机程序产品。应当理解,可使用基本上任何计算机架构而不背离本公开的范围。在不同的实施例中,计算设备 10 (图 1) 可以采取大型计算机、服务器计算机、台式计算机、膝上型计算机、平板计算机、家庭娱乐计算机、网络计算设备、移动计算设备、移动通信设备、游戏设备等等形式。

[0022] 计算系统 10 包括逻辑子系统 10a 和数据保持子系统 10b。计算系统 10 可以任选地包括显示子系统 10c、通信子系统(未示出)和/或在图 1 中未示出的其他组件。计算系统 10 还可以任选地包括诸如下列用户输入设备:例如键盘、鼠标、游戏控制器、相机、话筒和/或触摸屏等等。

[0023] 逻辑子系统 10a 包括被配置成执行一个或多个指令的一个或多个物理设备,例如包含在数据保持子系统 10b 中的被配置成如在图 2 的示例性方法所述确定评级并计算密切关系的指令。例如,逻辑子系统可被配置为执行一个或多个指令,该一个或多个指令是一个或多个应用程序、服务、程序、例程、库、对象、组件、数据结构、或其它逻辑构造的部分。可实现此类指令以执行任务、实现数据类型、变换一个或多个设备的状态、或以其它方式得到所需结果。

[0024] 数据保持子系统 10b 可包括一个或多个物理、非瞬时设备,这些设备被配置成保持数据和/或可由该逻辑子系统执行的指令,以实现此处描述的方法和过程。在实现这样的方法和过程时,可以变换数据保持子系统 10b 的状态(例如,以保持不同数据)。

[0025] 数据保持子系统 10b 可以包括可移动介质和/或内置设备。数据保持子系统 10b 可包括光学存储设备(例如,CD、DVD、HD-DVD、蓝光光盘等)、半导体存储器设备(例如,RAM、EPROM、EEPROM 等)、和/或磁存储器设备(例如,硬盘驱动器、软盘驱动器、磁带驱动器、MRAM 等)等。数据保持子系统 10b 可以包括具有以下特性中的一个或多个特性的设备:易



失性、非易失性、动态、静态、读 / 写、只读、随机存取、顺序存取、位置可寻址、文件可寻址、以及内容可寻址。在某些实施例中,可以将逻辑子系统 10a 和数据保持子系统 10b 集成到一个或多个常见设备中,如专用集成电路或片上系统。

[0026] 术语“模块”、“程序”和“引擎”可用于描述被实现以执行一个或更多个具体功能的计算系统 10 的一方面。在某些情况下,可以通过执行由数据保持子系统 10b 所保持的指令的逻辑子系统 10a 来实例化这样的模块、程序或引擎。应当理解,可以从同一应用、服务、代码块、对象、库、例程、API、函数等来实例化不同的模块、程序和 / 或引擎。同样,可以由不同的应用、服务、代码块、对象、例程、API、函数等来实例化同一模块、程序和 / 或引擎。术语“模块”、“程序”和“引擎”意在涵盖单个或成组的可执行文件、数据文件、库、驱动程序、脚本、数据库纪录等。

[0027] 当被包括时,显示子系统 10c 可用于呈现由数据保持子系统 10b 所保持的数据的可视表示。由于此处所描述的方法和过程改变了由数据保存子系统保存的数据,并由此转变了数据保存子系统的状态,因此同样可以转变显示子系统 10c 的状态以可视地表示底层数据中的改变。显示子系统 10c 可以包括使用实际上任何类型的技术的一个或多个显示设备。可将此类显示设备与逻辑子系统 10a 和 / 或数据保持子系统 10b 结合在共享外壳内,或此类显示设备可以是外围显示设备。

[0028] 当被包括时,通信子系统可以被配置成与具有一个或多个其他计算设备,例如存储内容项和在线评论或允许对内容项和在线评论的访问的设备,的计算系统 10 通信耦合。通信子系统可包括与一个或多个不同的通信协议相兼容的有线和 / 或无线通信设备。作为非限制性示例,该通信子系统可以被配置成经由无线电话网、无线局域网、有线局域网、无线广域网、有线广域网等进行通信。在一些实施例中,该通信子系统可允许计算系统 10 经由网络(比如因特网)向其他设备发送消息和 / 或从其他设备接收消息。

[0029] 应该理解,此处所述的配置和 / 或方法在本质上是示例性的,且这些具体实施例或示例不是局限性的,因为多个变体是可能。还有,在此的示例系统和方法不能被解释为是对解决特定问题或难题的方法和系统的限制。此处所述的具体例程或方法可表示任何数量的处理策略中的一个或更多个。由此,所示出的各个动作可以按所示顺序执行、按其他顺序执行、并行地执行、或者在某些情况下省略。同样,可以改变上述过程的次序。

[0030] 本发明的主题包括各种过程、系统和配置的所有新颖和非显而易见的组合和子组合、和此处所公开的其他特征、功能、动作、和 / 或特性、以及其任何和全部等效物。

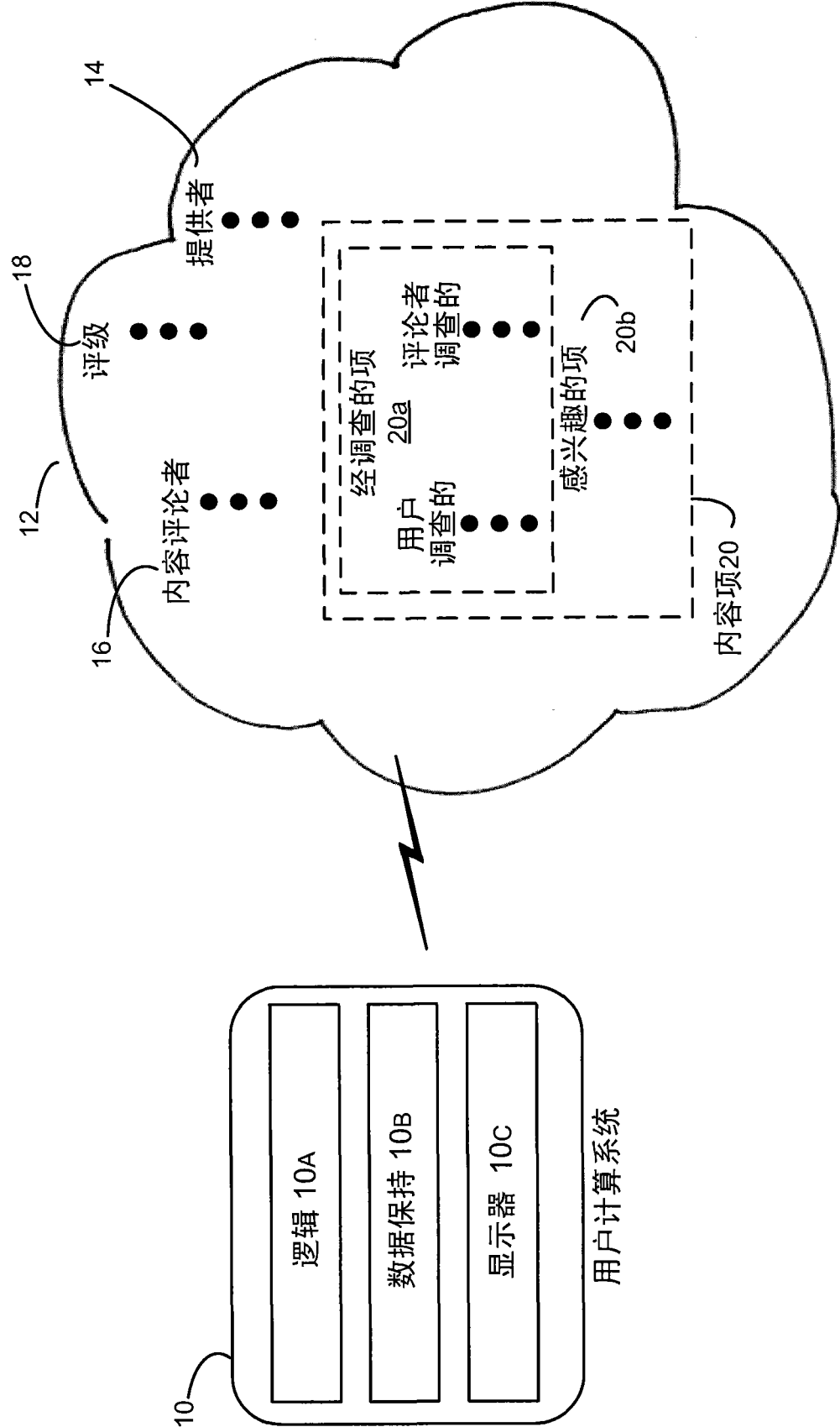


图 1

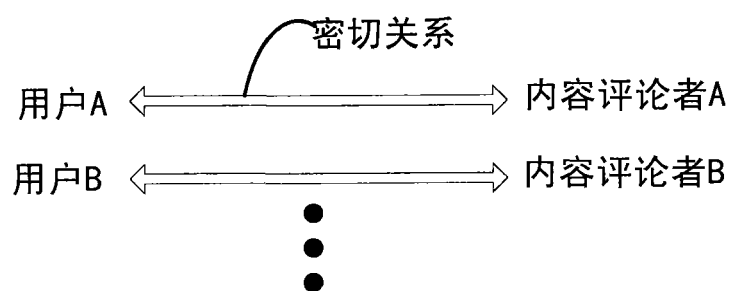


图 2

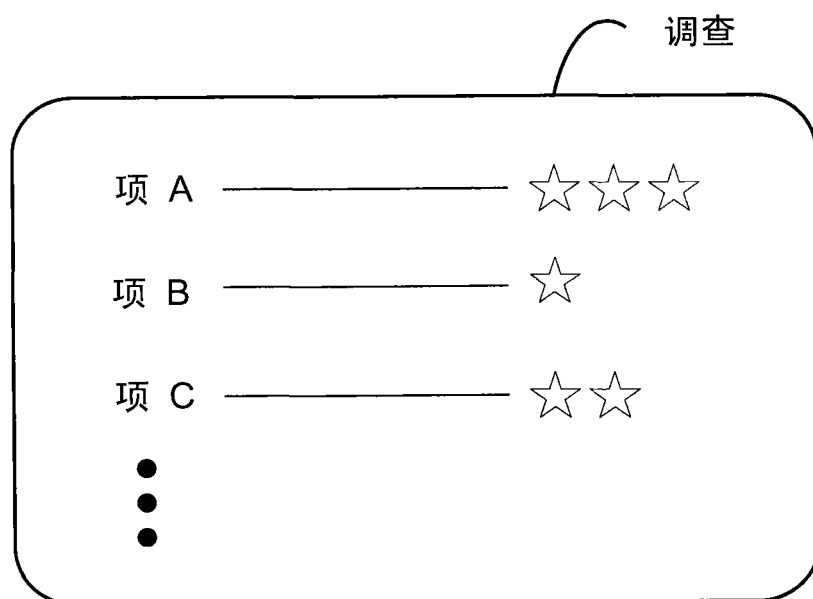


图 3

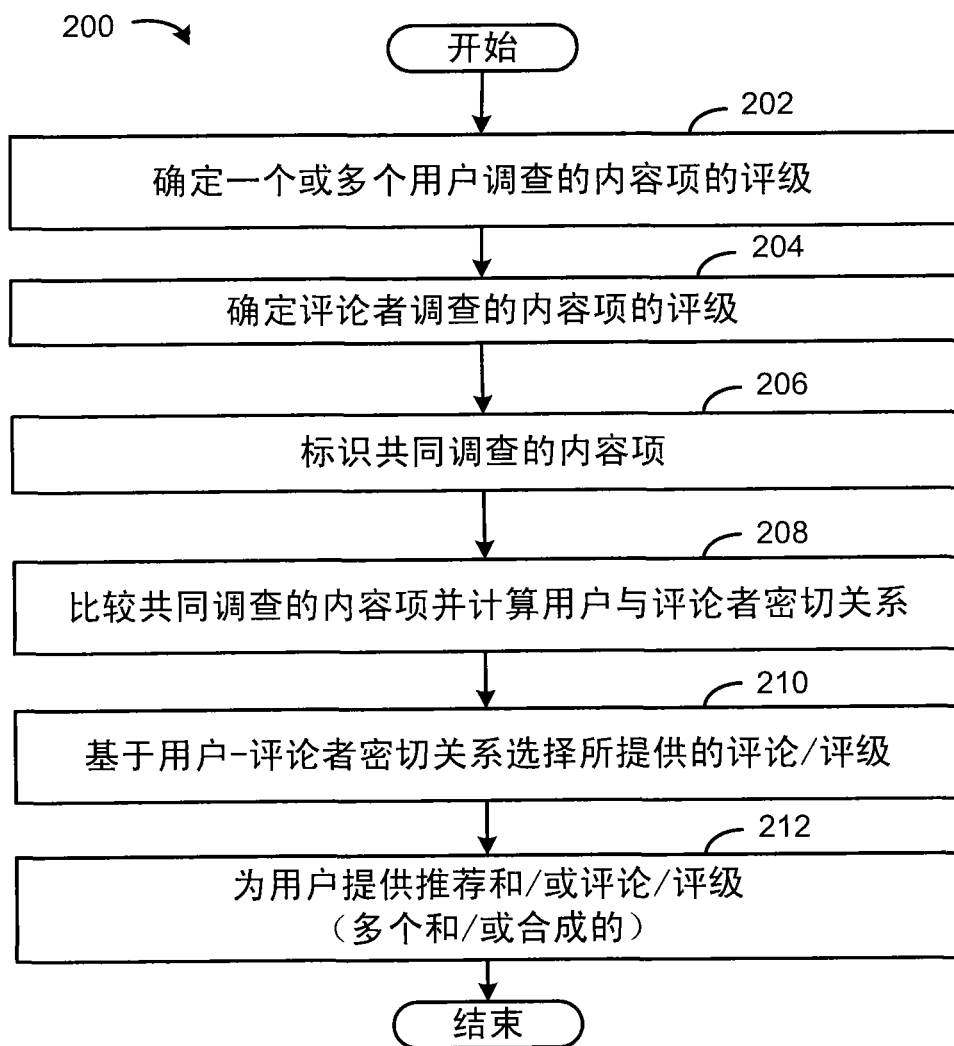


图 4