

(12) 发明专利申请

(10) 申请公布号 CN 102681999 A

(43) 申请公布日 2012. 09. 19

(21) 申请号 201110056046. 1

(22) 申请日 2011. 03. 08

(71) 申请人 阿里巴巴集团控股有限公司

地址 英属开曼群岛大开曼岛资本大厦一座  
四层 847 号邮箱

(72) 发明人 苏宁军 张旭 龙荣深 杨志雄

(74) 专利代理机构 北京同达信恒知识产权代理  
有限公司 11291

代理人 郭润湘

(51) Int. Cl.

G06F 17/30 (2006. 01)

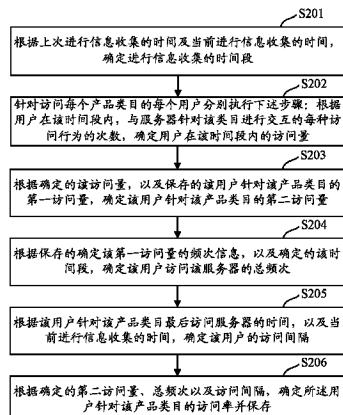
权利要求书 4 页 说明书 11 页 附图 4 页

(54) 发明名称

一种用户行为信息收集及信息发送方法及装  
置

(57) 摘要

本申请公开了一种用户行为信息收集及信息发送方法及装置,用以解决现有数据库的存储空间受限,导致信息推送不准确的问题。该方法根据用户在一段时间内针对每个产品类目,与服务器进行交互的每种访问行为的次数,确定该用户在该段时间内的访问量,并根据保存的该用户针对该产品类目的第一访问量,确定该用户针对该产品类目的第二访问量,并可以确定用户访问该服务器的总频次,以及访问间隔,从而可以确定用户针对该产品类目的长期偏好。如本申请提出的方案可以确定用户的长期偏好,保证向用户提供的信息的准确性。另外数据库无需一一保存每个用户的历史数据,减轻了数据库的存储压力,因此提高了数据库的工作效率。



1. 一种用户行为信息收集方法,其特征在于,包括:

根据上次进行信息收集的时间及当前进行信息收集的时间,确定进行信息收集的时间段;

在该时间段内,针对访问产品类目的用户分别执行下述步骤:

根据该用户在该时间段内,与服务器针对该产品类目进行交互的访问行为的次数,确定所述用户在该时间段内的访问量;

根据确定的该访问量,以及保存的该用户针对该产品类目的第一访问量,确定该用户针对该产品类目的第二访问量;

根据保存的确定该第一访问量的频次,以及确定的该时间段对应的频次,确定该用户访问该服务器的总频次;

根据该用户针对该产品类目最后访问服务器的时间,以及当前进行信息收集的时间,确定该用户的访问间隔;

根据确定的第二访问量、总频次以及访问间隔,确定所述用户针对该产品类目的长期偏好并保存。

2. 如权利要求 1 所述的方法,其特征在于,所述方法还包括:

采用确定的所述第二访问量,对所述第一访问量进行更新;

采用所述总频次,对保存的确定该第一访问量的频次信息进行更新。

3. 如权利要求 1 所述的方法,其特征在于,所述根据确定的第二访问量、总频次以及访问间隔,确定所述用户针对该产品类目的长期偏好包括:

确定该第二访问量与总频次的乘积,根据该乘积与该访问间隔的商,确定所述用户针对该产品类目的长期偏好。

4. 如权利要求 1 所述的方法,其特征在于,确定所述用户在该时间段内的访问量包括:

根据该用户在该时间段内与服务器针对该类目进行交互的访问行为的次数,以及访问行为对应的权重值,确定所述用户在该时间段内的访问量。

5. 如权利要求 1 所述的方法,其特征在于,所述方法还包括:

针对用户,根据确定的该用户针对产品类目的长期偏好,以及预设的产品类目数量阈值,选择长期偏好较大的该数量阈值对应数量的产品类目,保存该用户针对选择的该每个产品类目的长期偏好。

6. 一种基于权利要求 1 的收集方法的信息发送方法,其特征在于,所述方法包括:

根据接收到的所述用户登录服务器的信息,及数据库中保存的长期偏好,及短期偏好,确定是否保存有该用户的长期偏好和短期偏好中的至少一种;

当存在该用户的长期偏好和短期偏好中的至少一种时,根据该长期偏好和短期偏好中的至少一种对应的产品类目,将该产品类目的信息推送给所述用户。

7. 如权利要求 6 所述的方法,其特征在于,所述确定是否保存有该用户的长期偏好和短期偏好中的至少一种包括:

根据所述用户登录服务器的信息,及所述服务器生成的日志,或所述用户所在客户端保存的 Cookie 文件或 Flash 文件,获取所述用户当前的访问数据信息;

根据所述当前的访问数据信息,确定所述用户针对产品类目的当前偏好;

确定是否保存有所述用户的长期偏好、短期偏好和当前偏好中的至少一种。

8. 如权利要求 7 所述的方法,其特征在于,当存在该用户的长期偏好、短期偏好和当前偏好时,将该产品类目的信息推送给所述用户包括:

根据用户的长期偏好对应的产品类目,确定第一数量的产品类目信息;

根据用户的短期偏好对应的产品类目,确定第二数量的产品类目信息;

根据用户当前偏好对应的产品类目,确定第三数量的产品类目信息;

将确定的第一数量的产品类目、第二数量的产品类目以及第三数量的产品类目对应的信息推送给所述用户。

9. 如权利要求 7 所述的方法,其特征在于,当存在该用户的长期偏好、短期偏好和当前偏好时,将该产品类目的信息推送给所述用户包括:

根据用户的长期偏好、短期偏好和当前偏好对应的产品类目的交集,确定第四数量的产品类目信息;

根据用户的长期偏好、短期偏好和当前偏好中每两个偏好对应的产品类目的交集,确定第五数量的产品类目信息;

根据用户的长期偏好、短期偏好或当前偏好对应的每个产品类目,确定第六数量的产品类目信息;

将确定的第四数量的产品类目、第五数量的产品类目以及第六数量的产品类目对应的信息推送给所述用户。

10. 如权利要求 7 所述的方法,其特征在于,当存在该用户的长期偏好、短期偏好和当前偏好时,将该产品类目的信息推送给所述用户包括:

判断保存的所述用户访问服务器的总频次,是否大于设置的频次阈值;

当判断结果为是时,根据所述用户的短期偏好以及当前偏好对应的产品类目信息,向所述用户推荐相应产品类目的信息;

否则,根据所述用户的长期偏好及当前偏好对应的产品类目信息,向所述用户推荐相应产品类目的信息。

11. 如权利要求 7 所述的方法,其特征在于,当存在该用户的长期偏好、短期偏好和当前偏好时,将该产品类目的信息推送给所述用户包括:

根据保存的每个用户的类型,判断所述用户是否为商业用户;

当判断结果为是时,根据所述用户的长期偏好及当前偏好对应的产品类目信息,向所述用户推荐相应产品类目的信息;

否则,根据所述用户的短期偏好以及当前偏好对应的产品类目信息,向所述用户推荐相应产品类目的信息。

12. 一种用户行为信息收集装置,其特征在于,所述装置包括:

时间段确定模块,用于根据上次进行信息收集的时间及当前进行信息收集的时间,确定进行信息收集的时间段;

访问量确定模块,用于在该时间段内,针对访问产品类目的用户分别执行下述步骤:根据该用户在该时间段内,与服务器针对该产品类目进行交互访问行为的次数,确定所述用户在该时间段内的访问量;根据确定的该访问量,以及保存的该用户针对该产品类目的第一访问量,确定该用户针对该产品类目的第二访问量;

频次确定模块,用于根据保存的确定该第一访问量的频次,以及确定的该时间段对应

的频次,确定该用户访问该服务器的总频次;

时间间隔确定模块,用于根据该用户针对该产品类目最后访问服务器的时间,以及当前进行信息收集的时间,确定该用户的访问间隔;

偏好确定模块,用于根据确定的第二访问量、总频次以及访问间隔,确定所述用户针对该产品类目的长期偏好并保存。

13. 如权利要求 12 所述的装置,其特征在于,所述装置还包括:

更新模块,用于采用确定的所述第二访问量,对所述第一访问量进行更新;采用所述总频次,对保存的确定该第一访问量的频次信息进行更新。

14. 如权利要求 12 所述的装置,其特征在于,所述偏好确定模块,具体用于确定该第二访问量与总频次的乘积,根据该乘积与该访问间隔的商,确定所述用户针对该产品类目的长期偏好。

15. 如权利要求 12 所述的装置,其特征在于,所述访问量确定模块,具体用于根据该用户在该时间段内与服务器针对该类目进行交互的访问行为的次数,以及访问行为对应的权重值,确定所述用户在该时间段内的访问量。

16. 如权利要求 12 所述的装置,其特征在于,所述装置还包括:

过滤模块,用于针对用户,根据确定的该用户针对产品类目的长期偏好,以及预设的产品类目数量阈值,选择长期偏好较大的该数量阈值对应数量的产品类目,保存该用户针对选择的该产品类目的长期偏好。

17. 一种基于权利要求 12 的收集装置的信息发送装置,其特征在于,所述装置包括:

确定模块,用于根据接收到的所述用户登录服务器的信息,及数据库中保存的长期偏好,及短期偏好,确定是否保存有该用户的长期偏好和短期偏好中的至少一种;

推送模块,用于当存在该用户的长期偏好和短期偏好中的至少一种时,根据该长期偏好和短期偏好中的至少一种对应的产品类目,将该产品类目的信息推送给所述用户。

18. 如权利要求 17 所述的装置,其特征在于,所述确定模块,具体用于根据所述用户登录服务器的信息,及所述服务器生成的日志,或所述用户所在客户端保存的 Cookie 文件或 Flash 文件,获取所述用户当前的访问数据信息;根据所述当前的访问数据信息,确定所述用户针对每种产品类目的当前偏好;确定是否保存有所述用户的长期偏好、短期偏好和当前偏好中的至少一种。

19. 如权利要求 18 所述的装置,其特征在于,所述推送模块,具体用于当存在该用户的长期偏好、短期偏好和当前偏好时,根据用户的长期偏好对应的每个产品类目,确定第一数量的产品类目信息;根据用户的短期偏好对应的每个产品类目,确定第二数量的产品类目信息;根据用户当前偏好对应的每个产品类目,确定第三数量的产品类目信息;将确定的第一数量的产品类目、第二数量的产品类目以及第三数量的产品类目对应的信息推送给所述用户。

20. 如权利要求 18 所述的装置,其特征在于,所述推送模块,具体用于当存在该用户的长期偏好、短期偏好和当前偏好时,根据用户的长期偏好、短期偏好和当前偏好对应的产品类目的交集,确定第四数量的产品类目信息;根据用户的长期偏好、短期偏好和当前偏好中每两个偏好对应的产品类目的交集,确定第五数量的产品类目信息;根据用户的长期偏好、短期偏好或当前偏好对应的产品类目,确定第六数量的产品类目信息;将确定的第四

数量的产品类目、第五数量的产品类目以及第六数量的产品类目对应的信息推送给所述用户。

21. 如权利要求 18 所述的装置，其特征在于，所述推送模块，具体用于当存在该用户的长期偏好、短期偏好和当前偏好时，判断保存的所述用户访问服务器的总频次，是否大于设置的频次阈值；当判断结果为是时，根据所述用户的短期偏好以及当前偏好对应的产品类目信息，向所述用户推荐相应产品类目的信息；否则，根据所述用户的长期偏好及当前偏好对应的产品类目信息，向所述用户推荐相应产品类目的信息。

22. 如权利要求 18 所述的装置，其特征在于，所述推送模块，具体用于当存在该用户的长期偏好、短期偏好和当前偏好时，根据保存的每个用户的类型，判断所述用户是否为商业用户；当判断结果为是时，根据所述用户的长期偏好及当前偏好对应的产品类目信息，向所述用户推荐相应产品类目的信息；否则，根据所述用户的短期偏好以及当前偏好对应的产品类目信息，向所述用户推荐相应产品类目的信息。

## 一种用户行为信息收集及信息发送方法及装置

### 技术领域

[0001] 本申请涉及网络技术领域，尤其涉及一种用户行为信息收集及信息发送方法及装置。

### 背景技术

[0002] 当前服务器在向客户端推送信息时，一般都是基于数据库在设定时间长度内保存的用户行为信息，确定用户在短期内的兴趣偏好，从而可以向用户推动相应的信息，该设定的时间长度一般为一个月。

[0003] 现有技术中，数据库中一般只存储用户短期（设定时间长度）的行为信息，这样可以节省数据库的存储空间，但基于数据库中存储的该短期的用户行为信息，服务器只能确定用户在该设定时间长度内的兴趣偏好。当数据库中未保存用户在该设定时间长度内的行为信息时，服务器则无法确定该用户的兴趣偏好，并会将该用户确定为新用户，将按照新用户对应的信息推送类型，向该用户推送信息。但是实际上该用户可能在之前访问过数据库，例如该用户为周期性用户，会周期性的访问服务器。因此，现有技术中仅存储用户短期的行为信息虽然节省了数据库的存储空间，但由于存储的用户行为信息的数据量较少，从而导致服务器无法准确的确定该用户的兴趣偏好，影响向该用户推送的信息的准确性。如果增大现有技术中设定时间长度的跨度，虽然一定程度上提高了确定用户兴趣偏好的准确度，提高了向用户推送信息的准确度，但由于增大了数据库中存储的用户行为信息量，使得必须扩充数据库的存储空间，增加了硬件成本。

[0004] 现有技术中存在上述问题，主要是因为当时间跨度比较大时，用户的历史访问数据量非常的大，而现有数据库的存储空间有限，数据库中不可能长期保存用户的历史访问数据，从而无法确定用户在较长时间的兴趣偏好，也就影响了推送给用户的信息的准确性。

### 发明内容

[0005] 有鉴于此，本申请实施例提供一种用户行为信息收集及信息发送方法及装置，用以解决现有数据库的存储空间受限，导致信息推送不准确的问题。

[0006] 本申请实施例提供的一种用户行为信息收集方法，包括：

[0007] 根据上次进行信息收集的时间及当前进行信息收集的时间，确定进行信息收集的时间段；

[0008] 在该时间段内，针对访问产品类目的用户分别执行下述步骤：

[0009] 根据该用户在该时间段内，与服务器针对该产品类目进行交互的访问行为的次数，确定所述用户在该时间段内的访问量；

[0010] 根据确定的该访问量，以及保存的该用户针对该产品类目的第一访问量，确定该用户针对该产品类目的第二访问量；

[0011] 根据保存的确定该第一访问量的频次，以及确定的该时间段对应的频次，确定该用户访问该服务器的总频次；

- [0012] 根据该用户针对该产品类目最后访问服务器的时间,以及当前进行信息收集的时间,确定该用户的访问间隔;
- [0013] 根据确定的第二访问量、总频次以及访问间隔,确定所述用户针对该产品类目的长期偏好并保存。
- [0014] 本申请实施例提供的一种基于上述信息收集方法的信息发送方法,包括:
- [0015] 根据接收到的所述用户登录服务器的信息,及数据库中保存的长期偏好,及短期偏好,确定是否保存有该用户的长期偏好和短期偏好中的至少一种;
- [0016] 当存在该用户的长期偏好和短期偏好中的至少一种时,根据该长期偏好和短期偏好中的至少一种对应的产品类目,将该产品类目的信息推送给所述用户。
- [0017] 本申请实施例提供的一种用户行为信息收集装置,包括:
- [0018] 时间段确定模块,用于根据上次进行信息收集的时间及当前进行信息收集的时间,确定进行信息收集的时间段;
- [0019] 访问量确定模块,用于在该时间段内,针对访问产品类目的用户分别执行下述步骤:根据该用户在该时间段内,与服务器针对该产品类目进行交互的访问行为的次数,确定所述用户在该时间段内的访问量;根据确定的该访问量,以及保存的该用户针对该产品类目的第一访问量,确定该用户针对该产品类目的第二访问量;
- [0020] 频次确定模块,用于根据保存的确定该第一访问量的频次,以及确定的该时间段对应的频次,确定该用户访问该服务器的总频次;
- [0021] 时间间隔确定模块,用于根据该用户针对该产品类目最后访问服务器的时间,以及当前进行信息收集的时间,确定该用户的访问间隔;
- [0022] 偏好确定模块,用于根据确定的第二访问量、总频次以及访问间隔,确定所述用户针对该产品类目的长期偏好并保存。
- [0023] 本申请实施例提供的一种基于上述信息收集装置的信息发送装置,包括:
- [0024] 确定模块,用于根据接收到的所述用户登录服务器的信息,及数据库中保存的长期偏好,及短期偏好,确定是否保存有该用户的长期偏好和短期偏好中的至少一种;
- [0025] 推送模块,用于当存在该用户的长期偏好和短期偏好中的至少一种时,根据该长期偏好和短期偏好中的至少一种对应的产品类目,将该产品类目的信息推送给所述用户。
- [0026] 本申请实施例提供了一种用户行为信息收集及信息发送方法及装置,该信息收集方法根据用户在一段时间内针对产品类目,与服务器进行交互的访问行为的次数,确定该用户在该段时间内的访问量,并根据保存的该用户针对该产品类目的第一访问量,确定该用户针对该产品类目的第二访问量,并可以确定用户访问该服务器的总频次,以及访问间隔,从而可以确定用户针对该产品类目的长期偏好。由于在本申请实施例中通过保存用户的针对每种产品类目的第一访问量,以及用户在一段时间内的访问量,从而可以确定用户的第二访问量,也就是用户的总访问量,进而可以确定用户的长期偏好,保证向用户提供的信息的准确性。另外在本申请中数据库无需一一保存每个用户的历史数据,从而减轻了数据库的存储压力,由于数据库无需向服务器提供其所需的历史数据,因此提高了数据库的工作效率。

## 附图说明

- [0027] 图 1 为本申请实施例提供的一种用户信息收集系统的结构示意图；
- [0028] 图 2 为本申请实施例提供的一种用户行为信息收集过程；
- [0029] 图 3 为本申请实施例提供的一种基于上述信息收集方法的信息发送过程；
- [0030] 图 4 为本申请实施例提供的一种用户行为信息收集装置的结构示意图；
- [0031] 图 5 为本申请实施例提供的基于上述图 4 所示的装置的信息发送装置结构示意图。

## 具体实施方式

[0032] 本申请实施例提供的用户信息收集方法，可以确定用户的长期偏好，从而提高服务器提供给用户的信息的准确性。另外，由于在本申请实施例中数据库无需一一保存每个用户的历史数据，从而减轻了数据库的存储压力，由于数据库无需向服务器提供其所需的每个用户的历史数据，因此提高了数据库的工作效率。

[0033] 下面结合说明书附图，对本发明实施例进行详细说明。

[0034] 图 1 为本申请实施例提供的一种用户信息收集系统的结构示意图，该系统包括：服务器、数据库和客户端，其中，

[0035] 客户端将用户进行交互的每种访问行为发送到服务器。

[0036] 服务器接收到客户端发送的用户与其进行交互的每种访问行为时，根据该用户信息、产品类目信息、访问行为发生的时间信息以及该访问行为的信息，生成工作日志，并将该工作日志发送到数据库中保存；并，在进行用户行为信息收集时，根据上次进行信息收集的时间，以及当前进行信息收集的时间，确定进行信息收集的时间段。

[0037] 当服务器确定了进行信息收集的时间段后，根据数据库中保存的日志信息，由于该日志信息中记录有访问行为发生的时间信息，因此可以查找在该时间段内用户与服务器进行交互的每种访问行为。具体实施过程中，本申请所述的数据库可以与所述服务器集成在一台服务器中，也可以是独立于所述服务器而单独存在的数据库服务器。所述服务器可以是一台服务器，也可以是多台服务器组成的服务器集群。本申请对此并不作限定。

[0038] 具体的，由于服务器在向用户所在的客户端发送相应的信息时，为了保证发送的信息的准确性，服务器需要根据用户对每种产品类目的偏好的高低，向该用户推送偏好较高的产品类目的信息。并且，在本申请实施例中为了体现用户在长时间内，对某一产品类目信息的访问程度，可以采用该用户对该产品类目的长期偏好来表示。

[0039] 图 2 为本申请实施例提供的一种用户行为信息收集过程，该过程包括以下其步骤：

[0040] S201：根据上次进行信息收集的时间及当前进行信息收集的时间，确定进行信息收集的时间段。

[0041] 服务器可以根据设定的信息收集周期，定期进行用户行为信息的收集，或者也可以按照事件触发的条件，当满足某一事件的条件触发时，进行用户行为信息的收集，或者可以根据管理员的指示，进行用户行为信息的收集。

[0042] S202：在该时间段内，针对访问每个产品类目的每个用户分别执行下述步骤：根据该用户与服务器针对该类目进行交互的每种访问行为的次数，确定所述用户在该时间段内的访问量。

[0043] 由于数据库中保存有工作日志,该工作日志中包括用户信息、产品类目信息、访问行为发生的时间信息以及访问行为的信息等内容,因此服务器针对每个用户,根据该用户对每种产品类目进行交互的每种访问行为,确定该用户在该时间段内的访问量。

[0044] S203 :根据确定的该访问量,以及保存的该用户针对该产品类目的第一访问量,确定该用户针对该产品类目的第二访问量。

[0045] 其中为了提高服务器确定用户针对该产品类目的长期偏好的效率,该第一访问量可以保存在服务器中,当然为了节省服务器的存储空间,该第一访问量也可以保存在数据库服务器中,或其他网络设备中,当服务器对该用户针对该产品类目的长期偏好进行计算时,可以与数据库服务器或其他网络设备进行交互,获取该用户的针对该产品类目的第一访问量。

[0046] 在本申请实施例中,为了便于服务器确定每个用户针对每种产品类目的长期偏好,即用户在较长的时间长度内对某一产品类目的喜好程度,用户对某种产品类目的长期偏好可以通过用户在较长时间长度内对该产品类目的访问率来体现。在该服务器中需要该每个用户针对每种产品类目的第一访问量,即每个用户针对每种产品类目在上次进行信息收集结束后的访问量。根据该第一访问量,以及在该时间段内该用户针对每种产品类目的访问量,可以确定该用户针对每种产品类目在当前进行信息收集结束后的第二访问量。

[0047] 并且在本申请中当确定了用户针对某种产品类目的第二访问量后,由于该第二访问量为该用户到当前进行信息收集的时间为止,对该产品类目的访问量,因此为了便于下次对该用户针对该产品类目的长期偏好进行计算,采用该第二访问量对该第一访问量进行更新。

[0048] S204 :根据保存的确定该第一访问量的频次信息,以及确定的该时间段,确定该用户访问该服务器的总频次。

[0049] 为了便于服务器确定每个用户针对每种产品类目的长期偏好,本申请实施例中在该服务器中需要保存确定该第一访问量的频次信息,一般该频次信息可以采用天数来表示,具体的无论用户在一天中是否访问服务器,以及在一天中访问服务器多少次,都将该天累加到频次中作为一天。该第一访问量是根据用户首次针对该产品类目与服务器进行交互的当天,到上次进行信息收集的时间之间的天数中,用户对产品类目的访问量确定的,其中该总频次即为从用户针对该产品类目与服务器进行交互当天开始,到当前进行信息收集的时间之间的总的天数。

[0050] 例如用户首次对服务器的某一产品类目的访问时间为 2010. 3. 21,当前进行信息的时间为 2010. 4. 21,上次进行信息收集的时间为 2010. 3. 20,进行信息收集的时间段为 2010. 3. 21 ~ 2010. 4. 21,则针对当前进行信息收集的时间,由于之前该用户对该产品类目没有访问过,所以保存的该用户针对该产品类目的第一访问量为 0,该用户在 2010. 3. 21 ~ 2010. 4. 21 时间段内,与服务器针对该产品进行交互的访问行为的次数,可以从数据库中获取,从而可以确定用户在该时间段内的访问量,因此确定的第二访问量即为用户在该时间段内的访问量。由于之前用户针对该产品类目并未与服务器进行交互,因此确定该第一访问量的频次为 0,该时间段对应的天数为 31 天,因为可知该用户访问服务器的总频次为 31。服务器确定了该用户针对该产品类目的第二访问量以及总频次后,对自身保存的第一访问量,以及确定该第一访问量的频次进行更新。

[0051] 当用户下次再进行信息收集时,例如为 2010.5.21,以该时间为当前进行信息收集的时间,则上次进行信息收集的时间为 2010.4.21,进行信息收集的时间段为 2010.4.22 ~ 2010.5.21,根据在该时间段内用户与服务器针对该产品类目进行交互的访问行为的次数,可以确定用户在该时间段内的访问量,保存的第一访问量为用户首次对服务器的某一产品类目的访问,到上次进行信息收集时间范围内,用户对该产品类目的访问量,因此根据保存的第一访问量和确定的该用户在该时间段内的访问量,确定第二访问量,其中该第二访问量即为用户首次对服务器的某一产品类目的访问,到当前进行信息收集时间范围内,用户对该产品类目的访问量。保存的确定第一访问量的频次,即天数为 31 天,当时间段为 30 天,因此用户访问服务器的总频次为 61 天。之后继续根据确定的第二访问量以及总频次,对保存的第一访问量以及确定第一访问量的频次进行更新,并进行后续步骤,这里就不一一赘述。

[0052] S205 :根据该用户针对该产品类目最后访问服务器的时间,以及当前进行信息收集的时间,确定该用户的访问间隔。

[0053] 在本申请实施例中该访问间隔可以采用天数来标识。该访问间隔为用户针对该产品类目最后访问服务器的那天,与进行信息收集的当天的天数差。

[0054] S206 :根据确定的第二访问量、总频次以及访问间隔,确定所述用户针对该产品类目的偏好并保存。

[0055] 具体的根据确定的第二访问量、总频次以及访问间隔,确定所述用户针对该产品类目的偏好并保存,包括 :确定该第二访问量与总频次的乘积,根据该乘积与该访问间隔的商,确定所述用户针对该产品类目的长期偏好。

[0056] 在服务器中为了确定每个用户针对每种产品类目的长期偏好,在本申请实施例中服务器需要保存上次进行用户行为信息收集的时间。因此当服务器当前进行用户行为信息的收集时,根据当前进行信息收集的时间,以及保存的上次进行信息收集的时间,可以确定进行信息收集的时间段。例如上次进行信息收集的时间为 2011 年 1 月 1 日零晨,当前进行信息收集的时间为 2011 年 1 月 31 日零晨,则确定进行信息收集的时间段为 2011 年 1 月 1 日至 2011 年 1 月 30 日。

[0057] 此时,服务器根据数据库中保存的工作日志,对该工作日志进行解析,获取访问行为发生的时间位于该时间段内的工作日志。具体的在本申请实施例中该访问行为包括 :搜索行为、浏览行为、点击行为、反馈行为、交易行为等其中的一种或几个。

[0058] 服务器在确定每个用户在该时间段内的访问量时,针对每个用户,根据获取的访问行为发生时间位于该时间段内的工作日志,查找包含该用户信息的工作日志,在包含该用户信息的工作日志中,查找包含某一产品类目信息的工作日志,在包含该用户信息及该某一产品类目的工作日志中,查找该用户与服务器进行交互的每种访问行为的次数。

[0059] 例如,服务器根据用户 A 针对产品类目 B 与其进行交互的每种访问行为的次数,确定该用户在该时间段内的访问量为例进行说明。首先服务器在获取的访问行为发生时间位于该时间段的工作日志中,查找包含用户 A 及产品类目 B 信息的工作日志,在查找到的工作日志中,分别统计用户 A 进行交互的搜索行为、浏览行为、点击行为、反馈行为以及交易行为等的次数,例如分别为  $x_1, K, x_n$ , 其中 n 为访问行为包含的种类数。

[0060] 确定了该用户在该时间段内,与服务器针对该产品类目进行交互的每种访问行为

的次数后,需要确定该用户在该时间段内的访问量,具体在确定用户在该时间段内的访问量时,可以直接根据确定的每种访问行为的次数的和,确定该访问量。另外,也可以针对每种访问行为预设不同的权重值,具体的例如可以认为用户主动发送访问行为的权重值较大,即可以预设搜索行为、点击行为和交易行为的权重值较大等。当针对每种访问行为预设了不同的权重值后,可以根据该用户在该时间段内与服务器针对该类目进行交互的每种访问行为的次数,以及每种访问行为对应的权重值,确定所述用户在该时间段内的访问量。即根据  $Y = w_1x_1 + \dots + w_nx_n$  确定用户在该时间段内的访问量,其中 Y 为用户在该时间段内的访问量,  $x_1, \dots, x_n$  为 n 中访问行为的次数,  $w_1, \dots, w_n$  为每种访问行为对应的权重值。

[0061] 服务器根据获取的工作日志,确定了该用户在该时间段内的访问量后,还需要根据保存的该用户针对该产品类目的第一访问量,确定该用户针对该产品类目的第二访问量,即该用户针对该产品类目到当前进行信息收集的时间为止的总的访问量。

[0062] 服务器确定了该用户针对该产品类目的第二访问量后,还需要根据保存的确定该第一访问量的频次信息,以及确定的该时间段,确定该用户针对该产品类目访问服务器的总频次,即当前进行该用户针对该产品类目的长期偏好确定时,进行信息收集的总的频次。

[0063] 当服务器确定了该用户针对该产品类目的第二访问量,以及总的频次后,根据该用户针对该产品类目最后访问服务器的时间,以及当前进行信息收集的时间,确定该用户针对该产品类目的访问间隔,即可确定用户针对该产品类目的长期偏好并保存。具体的以 Y 表示该用户针对该产品类目的第二访问量, F 为总的频次、 T 为用户针对该产品类目的访问间隔,则该用户针对该产品类目的长期偏好  $P = Y*F/T$ 。

[0064] 依据上述方法,服务器可以根据数据库中记录的工作日志,确定每个用户针对每个产品类目的长期偏好并保存。由于访问数据的用户的数量非常的大,如果在数据库中保存每个用户访问每种产品类目的长期偏好的话,在服务器中占用的存储空间也是非常大的。在本发明实施例中为了减小保存长期偏好占用的服务器的存储空间,可以针对每个产品类目,预设用户数量阈值。当针对该产品类目,确定了每个用户针对该产品类目的长期偏好后,将确定的每个用户针对该产品类目的长期偏好进行排序,根据该产品类目对应的预设的用户数量阈值,选择长期偏好较大的该数量阈值对应数量的用户,保存该每个用户针对该产品类目的长期偏好。

[0065] 或者,也可以针对每种用户,根据确定的该用户针对每个产品类目的长期偏好,以及预设的产品类目数量阈值,选择长期偏好较大的该数量阈值对应数量的产品类目,保存该用户针对选择的该每个产品类目的长期偏好。

[0066] 服务器保存了每个用户针对每种产品类目的长期偏好后,为了便于服务器后期进行每个用户针对每种产品类目的长期偏好的确定,在本申请实施例中,服务器采用确定的该用户针对该产品类目的第二访问量,更新自身保存的该用户针对该产品类目的第一访问量,并采用该用户针对该产品类目访问该服务器的总频次,更新保存的确定该第一访问量的频次信息。并且本申请实施例中只要能够获取第一访问量,确定第一访问量的频次,用户针对产品类目最后访问服务器的时间,以及进行信息收集的时间段内的用户与服务器进行交互的日志,就可以确定用户针对该产品类目的长期偏好。因此用户在上次进行信息收集的时间之前的访问日志即可删除,只要保存用户上次进行信息收集的时间,以及当前进行信息收集的时间对应的时间段内用户与服务器进行交互的日志,以及第一访问量,确定第

一访问量的频次以及用户针对该产品类目最后访问服务器的时间信息即可，因此大大节省了服务器的存储资源。

[0067] 对于用户针对该产品类目的访问间隔的更新，在本申请实施例中由于该访问间隔为该用户针对该产品类目最后访问服务器的时间，以及当前进行信息收集的时间的差，当在该时间段内，该用户针对该产品类目未与服务器进行交互时，服务器直接根据自身保存的该用户针对该产品类目的第一访问间隔，以及当前进行信息收集的时间段，确定用户针对该产品类目的访问间隔，并采用确定的该时间间隔更新服务器中保存的，该用户针对该产品类目的访问间隔。

[0068] 即当该用户在该时间段内针对该产品类目未与服务器进行交互时，则该用户针对该产品类目最后访问服务器的时间不在该时间段内，例如进行信息收集的时间段为 2011 年 1 月 1 日至 2011 年 1 月 30 日，用户在该时间段内未针对某一产品类目与服务器进行交互，则可知该用户针对该产品类目最后访问服务器的时间不在该时间段，应该在 2011 年 1 月 1 日之前的时间段内。因此在该服务器中保存了该用户针对该产品类目的第一访问间隔，该第一时间间隔为该用户针对该产品类目最后一个访问服务器的时间，以及上次进行信息收集的时间确定，因此可知当前进行信息收集的时间内，该用户针对该访问类目的访问时间间隔为该第一访问间隔与该进行信息收集的时间段的和。

[0069] 当在该时间段内用户针对该产品与服务器进行交互时，则根据该用户针对该产品类目最后访问服务器的时间，以及当期进行信息收集的时间，确定该用户的访问时间间隔，并采用确定的该时间间隔更新服务器中保存的，该用户针对该产品类目的访问间隔。

[0070] 由于服务器中保存了每个用户针对每种产品类目的第一访问量，总频次，以及访问间隔，因此当服务器确定每个用户针对每种产品类目的长期偏好时，只需收集当前进行信息收集的时间以及上次进行信息收集的时间差对应的时间段内，记录的工作日志信息，即可确定用户针对每种产品类目的长期偏好，从而无需数据库中长期保存用户的历史数据，因此本申请实施例提供的确定用户针对每种产品类目的长期偏好的方法，有效的节省了数据库的内存空间，并且由于可以根据本申请实施例提供的方法确定用户针对每种产品类目的长期偏好，因此服务器在进行信息发送时，可以保证发送的信息的准确性。

[0071] 另外，由于现有确定用户针对每种产品类目的短期偏好时，都是根据用户在当前进行信息收集的时间之前的设定频次内，与服务器针对每种产品类目进行交互的行为信息确定的，该短期偏好可以反映用户短期内的访问习惯。

[0072] 在确定用户针对每种产品类目的短期偏好时，根据每天该用户针对该产品类目与服务器进行交互的每种访问行为的次数，确定每天该用户针对该产品类目的访问量  $Y_i$ 。并根据确定的随时间  $t$  衰减的模型  $P(t) = K_1 + \exp((t - K_2) / K_3)$  确定用户针对该产品类目的短期偏好，其中  $t$  为该设定频次内每天对应的负数，例如当该为该设定频次内的第 5 天时，则  $t$  为 -5，参数  $K_1, K_2, K_3$  可以根据具体的应用确定。当确定的用户针对该产品类目的访问量，以及预设的衰减模型后，而可以得到用户针对该产品类目的短期偏好  $P(0) Y_0 + P(N) Y_N$ 。

[0073] 另外，现有服务器在确定每个用户针对每种产品类目的偏好时，由于数据库中数据的更新时间粒度一般为到天的。因此当用户与服务器针对某种产品类目进行交互时，服务器只有在交互后的第二天才能从数据库中，获取记录该交互过程的工作日志，因此现有服务器无法根据用户当前针对某一产品类目进行的交互，生成用户针对该产品类目的当前

偏好。

[0074] 在本申请实施例中服务器为了生成用户针对每种产品类目的当前偏好,当用户登录服务器时,服务器根据当前用户针对某种产品类目进行交互的访问行为,生成工作日志,在将该工作日志发送到数据库之前,服务器解析获取该工作日中记录的该用户针对该产品类目的访问行为,获取所述用户当前的访问数据信息;根据所述当前的访问数据信息,确定所述用户针对每种产品类目的当前偏好。

[0075] 或者,由于用户所在的客户端会在本地将用户通过该客户端与服务器针对某产品类目交互的行为信息,记录在本地的Cookie文件或Flash文件。因此服务器在生成用户针对每种类目的当前偏好时,可以与客户端进行交互,获取用户所在客户端本地记录的Cookie文件或Flash文件,记录的用户当前的访问数据信息,根据获取的该用户当前的访问数据信息,确定所述用户针对每种产品类目的当前偏好。

[0076] 由于在本申请实施例中,服务器可以确定用户针对每种产品类目的长期偏好,短期偏好以及当前偏好,因此在向用户发送信息时,可以根据保存的偏好进行发送,从而保证发送的信息的准确性。

[0077] 图3为本申请实施例提供的一种基于上述信息收集方法的信息发送过程,该过程包括以下步骤:

[0078] S301:接收用户登录服务器的信息。

[0079] S302:根据数据库中保存的长期偏好,及短期偏好,确定是否保存有该用户的长期偏好和短期偏好中的至少一种,当判断结果为是时,进行步骤S303,否则,进行步骤S304。

[0080] S303:根据该长期偏好和短期偏好中的至少一种对应的产品类目,将该产品类目的信息推送给所述用户。

[0081] S304:将所述用户作为新用户,将新用户对应的产品类目信息发送给所述用户。

[0082] 在本申请实施例中由于服务器中保存了用户针对每种产品类目的长期偏好。短期偏好以及当前偏好。当服务器接收到用户的登录信息后,根据保存的该用户对应的每种产品类目的偏好,向该用户发送相应产品类目的信息。

[0083] 当服务器中保存有该用户针对每种产品类目的长期偏好时,则可以根据保存的该用户针对每种产品类目的长期偏好的大小,将长期偏好较大的产品类目的信息发送给所述用户。当服务器中保存有该用户针对每种产品类目的短期偏好时,则可以根据保存的该用户针对每种产品类目的短期偏好的大小,将短期偏好较大的产品类目的信息发送给所述用户。当服务器中保存有该用户针对每种产品类目的当前偏好时,则可以根据保存的该用户针对每种产品类目的当前偏好的大小,将当前偏好较大的产品类目的信息发送给所述用户。

[0084] 当服务器中保存有该该用户针对每种产品类目的长期偏好,短期偏好以及当前偏好时,在向该用户发送信息时,可以根据用户的长期偏好对应的每个产品类目,确定第一数量的产品类目信息;根据用户的短期偏好对应的每个产品类目,确定第二数量的产品类目信息;根据用户当前偏好对应的每个产品类目,确定第三数量的产品类目信息;将确定的第一数量的产品类目、第二数量的产品类目以及第三数量的产品类目对应的信息推送给所述用户。

[0085] 即根据用户针对每种产品类目的长期偏好,将用户针对每种产品类目的长期偏好

排序,选择长期偏好较大的第一数量 N1 的产品类目,并采用相同的方法,选择短期偏好较大的第二数量 N2 的产品类目,同样依据相同的方法,选择当前偏好较大的第三数量 N3 的产品类目,将第一数量的产品类目、第二数量的产品类目以及第三数量的产品类目对应的信息推送给所述用户。

[0086] 或者,当服务器中保存有该该用户针对每种产品类目的长期偏好,短期偏好以及当前偏好时,在向该用户发送信息时,根据用户的长期偏好、短期偏好和当前偏好对应的产品类目的交集,确定第四数量的产品类目信息;根据用户的长期偏好、短期偏好和当前偏好中每两个偏好对应的产品类目的交集,确定第五数量的产品类目信息;根据用户的长期偏好、短期偏好或当前偏好对应的每个产品类目,确定第六数量的产品类目信息;

[0087] 将确定的第四数量的产品类目、第五数量的产品类目以及第六数量的产品类目对应的信息推送给所述用户。

[0088] 即首先确定该用户针对哪些产品类目,即存在长期偏好、短期偏好,也存在当前偏好,当确定了这些产品类目后,选择该第四数量的该产品数目,之后,确定该用户针对哪些产品类目,存在长期间率、短期偏好和当前偏好中的两个,在这些产品类目中,选择第五数量的该产品数目,再之后,根据哪些只存在长期偏好,短期偏好或当前偏好中的一种的产品类目,选择第六数量的产品类目,将确定的第四数量的产品类目、第五数量的产品类目以及第六数量的产品类目对应的信息推送给所述用户。

[0089] 再或者,当服务器中保存有该该用户针对每种产品类目的长期偏好,短期偏好以及当前偏好时,也可以根据用户的活跃度,向用户发送相应产品类目的信息,即判断保存的所述用户访问服务器的总频次,是否大于设置的频次阈值;当判断结果为是时,根据所述用户的短期偏好以及当前偏好对应的产品类目信息,向所述用户推荐相应产品类目的信息;否则,根据所述用户的长期偏好及当前偏好对应的产品类目信息,向所述用户推荐相应产品类目的信息。

[0090] 或者,当服务器中保存有该该用户针对每种产品类目的长期偏好,短期偏好以及当前偏好时,也可以根据用户的类型,向用户发送相应产品类目的信息,即根据保存的每个用户的类型,判断所述用户是否为商业用户;当判断结果为是时,根据所述用户的长期偏好及当前偏好对应的产品类目信息,向所述用户推荐相应产品类目的信息;否则,根据所述用户的短期偏好以及当前偏好对应的产品类目信息,向所述用户推荐相应产品类目的信息。

[0091] 图 4 为本申请实施例提供的一种用户行为信息收集装置的结构示意图,该装置包括:

[0092] 时间段确定模块 41,用于根据上次进行信息收集的时间及当前进行信息收集的时间,确定进行信息收集的时间段;

[0093] 访问量确定模块 42,用于在该时间段内,针对访问每个产品类目的每个用户分别执行下述步骤:根据该用户在该时间段内,与服务器针对该产品类目进行交互的每种访问行为的次数,确定所述用户在该时间段内的访问量;根据确定的该访问量,以及保存的该用户针对该产品类目的第一访问量,确定该用户针对该产品类目的第二访问量;

[0094] 频次确定模块 43,用于根据保存的确定该第一访问量的频次,以及确定的该时间段对应的频次,确定该用户访问该服务器的总频次;

[0095] 时间间隔确定模块 44,用于根据该用户针对该产品类目最后访问服务器的时间,

以及当前进行信息收集的时间,确定该用户的访问间隔;

[0096] 偏好确定模块 45,用于根据确定的第二访问量、总频次以及访问间隔,确定所述用户针对该产品类目的长期偏好并保存。

[0097] 所述装置还包括:

[0098] 更新模块 46,用于采用确定的所述第二访问量,对所述第一访问量进行更新;采用所述总频次,对保存的确定该第一访问量的频次信息进行更新。

[0099] 所述偏好确定模块 45,具体用于确定该第二访问量与总频次的乘积,根据该乘积与该访问间隔的商,确定所述用户针对该产品类目的长期偏好。

[0100] 所述访问量确定模块 42,具体用于根据该用户在该时间段内与服务器针对该类目进行交互的每种访问行为的次数,以及每种访问行为对应的权重值,确定所述用户在该时间段内的访问量。

[0101] 所述装置还包括:

[0102] 过滤模块 47,用于针对每种用户,根据确定的该用户针对每个产品类目的长期偏好,以及预设的产品类目数量阈值,选择长期偏好较大的该数量阈值对应数量的产品类目,保存该用户针对选择的该每个产品类目的长期偏好。

[0103] 图 5 为本申请实施例提供的基于上述图 4 所示的装置的信息发送装置结构示意图,该装置包括:

[0104] 确定模块 51,用于根据接收到的所述用户登录服务器的信息,及数据库中保存的长期偏好,及短期偏好,确定是否保存有该用户的长期偏好和短期偏好中的至少一种;

[0105] 推送模块 52,用于当存在该用户的长期偏好和短期偏好中的至少一种时,根据该长期偏好和短期偏好中的至少一种对应的产品类目,将该产品类目的信息推送给所述用户。

[0106] 所述确定模块 51,具体用于根据所述用户登录服务器的信息,及所述服务器生成的日志,或所述用户所在客户端保存的 Cookie 文件或 Flash 文件,获取所述用户当前的访问数据信息;根据所述当前的访问数据信息,确定所述用户针对每种产品类目的当前偏好;确定是否保存有所述用户的长期偏好、短期偏好和当前偏好中的至少一种。

[0107] 所述推送模块 52,具体用于当存在该用户的长期偏好、短期偏好和当前偏好时,根据用户的长期偏好对应的每个产品类目,确定第一数量的产品类目信息;根据用户的短期偏好对应的每个产品类目,确定第二数量的产品类目信息;根据用户当前偏好对应的每个产品类目,确定第三数量的产品类目信息;将确定的第一数量的产品类目、第二数量的产品类目以及第三数量的产品类目对应的信息推送给所述用户。

[0108] 所述推送模块 52,具体用于当存在该用户的长期偏好、短期偏好和当前偏好时,根据用户的长期偏好、短期偏好和当前偏好对应的产品类目的交集,确定第四数量的产品类目信息;根据用户的长期偏好、短期偏好和当前偏好中每两个偏好对应的产品类目的交集,确定第五数量的产品类目信息;根据用户的长期偏好、短期偏好或当前偏好对应的产品类目,确定第六数量的产品类目信息;将确定的第四数量的产品类目、第五数量的产品类目以及第六数量的产品类目对应的信息推送给所述用户。

[0109] 所述推送模块 52,具体用于当存在该用户的长期偏好、短期偏好和当前偏好时,判断保存的所述用户访问服务器的总频次,是否大于设置的频次阈值;当判断结果为是时,根

据所述用户的短期偏好以及当前偏好对应的产品类目信息,向所述用户推荐相应产品类目的信息;否则,根据所述用户的长期偏好及当前偏好对应的产品类目信息,向所述用户推荐相应产品类目的信息。

[0110] 所述推送模块 52,具体用于当存在该用户的长期偏好、短期偏好和当前偏好时,根据保存的每个用户的类型,判断所述用户是否为商业用户;当判断结果为是时,根据所述用户的长期偏好及当前偏好对应的产品类目信息,向所述用户推荐相应产品类目的信息;否则,根据所述用户的短期偏好以及当前偏好对应的产品类目信息,向所述用户推荐相应产品类目的信息。

[0111] 本申请实施例提供了一种用户行为信息收集及信息发送方法及装置,该信息收集方法根据用户在一段时间内针对每个产片类目,与服务器进行交互的每种访问行为的次数,确定该用户在该段时间内的访问量,并根据保存的该用户针对该产品类目的第一访问量,确定该用户针对该产品类目的第二访问量,并可以确定用户访问该服务器的总频次,以及访问间隔,从而可以确定用户针对该产品类目的长期偏好。由于在本申请实施例中通过保存用户的针对每种产品类目的第一访问量,以及用户在一段时间内的访问量,从而可以确定用户的第二访问量,也就是用户的总访问量,进而可以确定用户的长期偏好,保证向用户提供信息的准确性。另外在本申请中数据库无需一一保存每个用户的历史数据,从而减轻了数据库的存储压力,由于数据库无需向服务器提供其所需的历史数据,因此提高了数据库的工作效率。

[0112] 显然,本领域的技术人员可以对本申请进行各种改动和变型而不脱离本申请的精神和范围。这样,倘若本申请的这些修改和变型属于本申请权利要求及其等同技术的范围之内,则本申请也意图包含这些改动和变型在内。

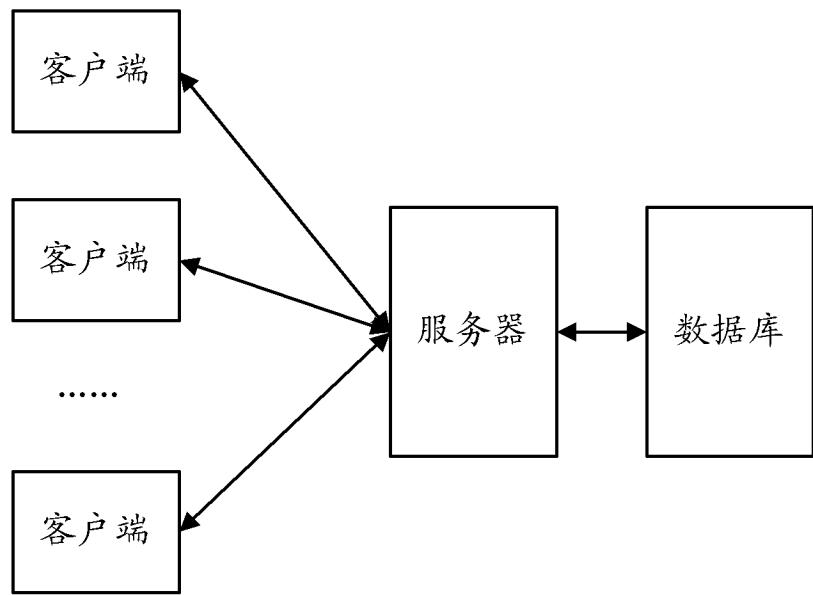


图 1

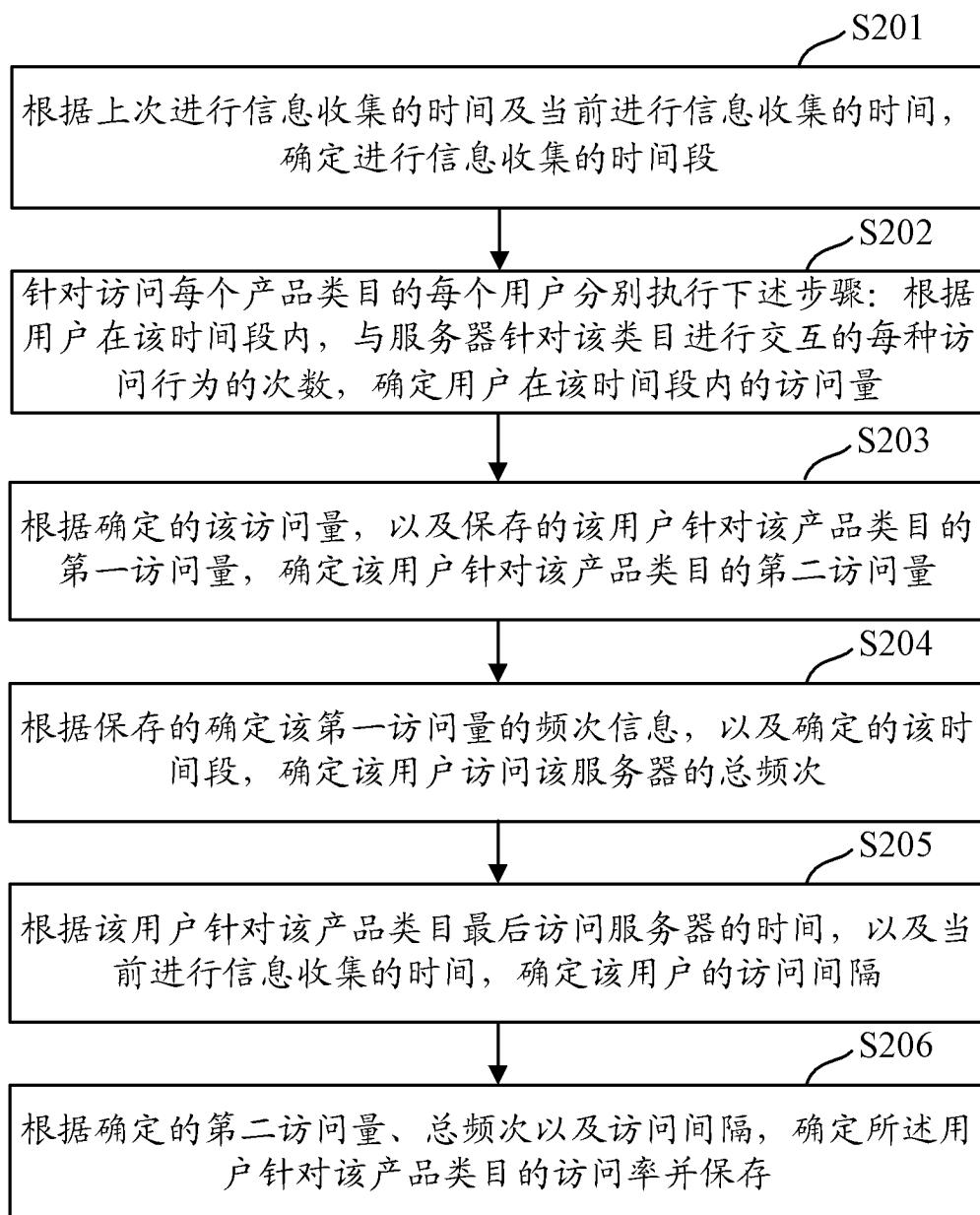


图 2

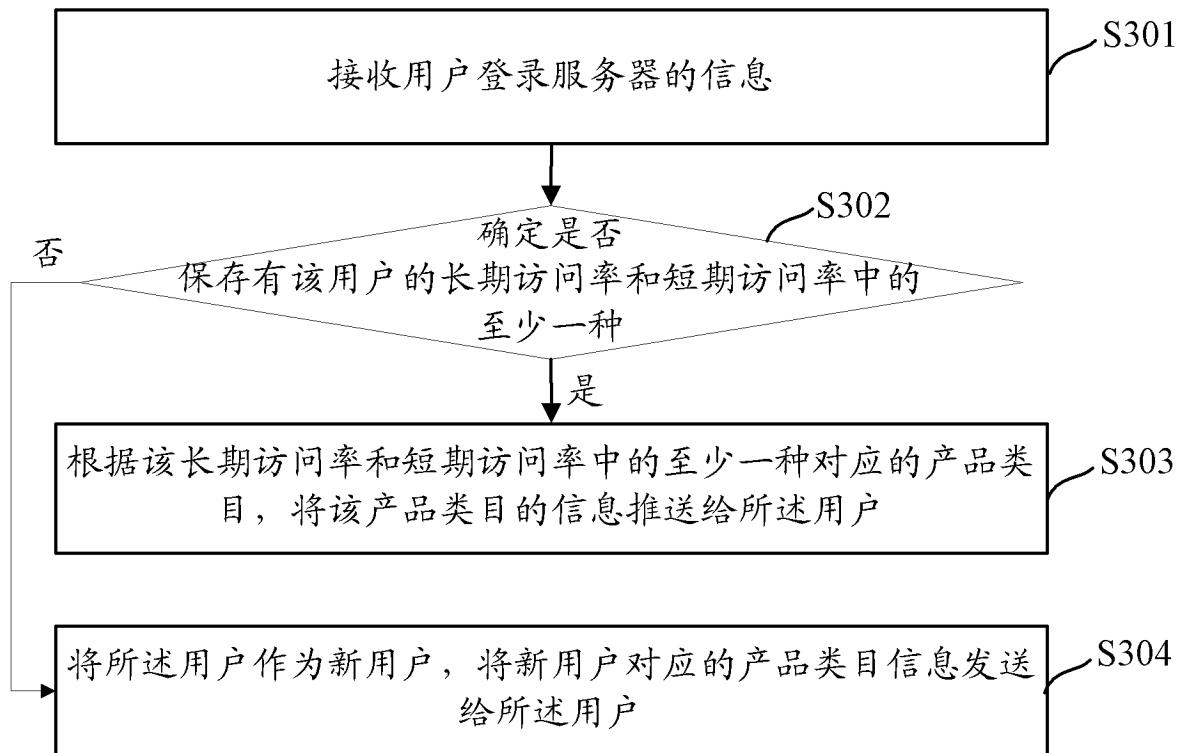


图 3

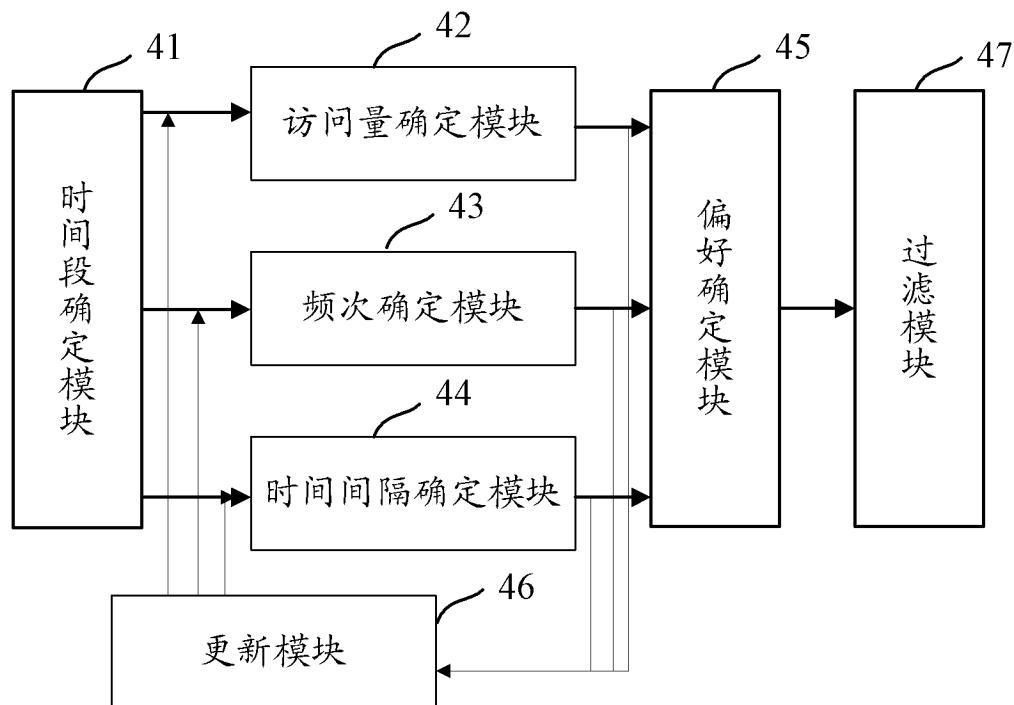


图 4

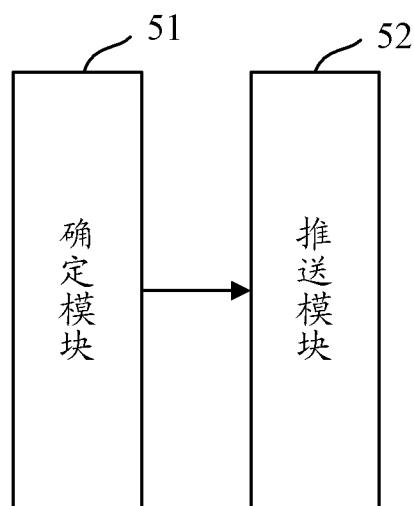


图 5