



(12) 发明专利

(10) 授权公告号 CN 102253936 B

(45) 授权公告日 2013. 07. 24

(21) 申请号 201010176800. 0

CN 1858733 A, 2006. 11. 08, 全文.

(22) 申请日 2010. 05. 18

WO 2007/106403 A2, 2007. 09. 20, 全文.

(73) 专利权人 阿里巴巴集团控股有限公司

审查员 邹予婷

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 袁伟

(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)

G06Q 30/02 (2012. 01)

(56) 对比文件

US 2006/0224406 A1, 2006. 10. 05, 说明书第
5 页第 76 段到第 15 页第 206 段、权利要求 1-30、
附图 1-41.

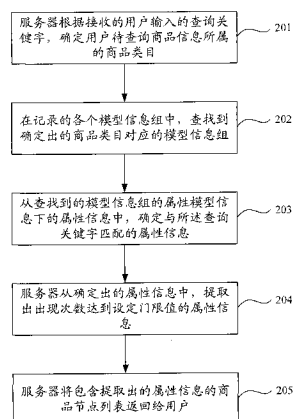
权利要求书2页 说明书10页 附图3页

(54) 发明名称

记录用户访问商品信息的方法及搜索方法和
服务器

(57) 摘要

本申请公开了一种记录用户访问商品信息的方法和服务器, 以及利用记录的用户访问商品信息的结果对商品信息进行搜索的方法和服务器, 主要内容包括: 对一段时间内用户对商品信息的访问进行记录, 根据记录的汇总数据确定用户对商品信息的查询意图分布情况, 得到用户对商品信息访问意图, 进而利用得到的用户对商品信息的访问意图对用户的搜索请求进行分类, 将与用户搜索的商品信息相关性较高的包含商品信息页面链接的商品信息节点列表返回给用户, 让搜索结果更加接近用户的真实意图, 使得搜索结果更好。



1. 一种记录用户访问商品信息的方法,其特征在于,所述方法包括:

服务器收到用户对包含商品信息的页面的访问信息时,生成日志文件,所述日志文件中包含访问页面中包含的商品信息的至少一个属性信息;

根据生成的多个所述日志文件中包含的属性信息,分别确定同一商品类目对应的各个不同属性信息及每种属性信息出现的次数信息;

将确定出的同一商品类目对应的各个不同属性信息以及每种属性信息出现的次数信息作为一模型信息组记录;其中:所述模型信息组包括属性模型信息,每一所述属性模型信息包括一属性信息及其对应出现的次数信息。

2. 如权利要求1所述的方法,其特征在于,所述日志文件中包含的商品信息的属性信息为:商品品牌信息、商品型号信息、商品颜色信息和商品类目信息;

所述模型信息组中包含以下属性模型信息:

由商品品牌信息及其出现次数构成的商品品牌模型信息;

由商品型号信息及其出现次数构成的商品型号模型信息;

由商品颜色信息及其出现次数构成的商品颜色模型信息。

3. 一种对商品信息进行搜索的方法,其特征在于,利用权利要求1记录的用户访问商品信息,所述方法包括:

服务器根据查询关键字,确定待查询商品信息所属的商品类目;

在记录的各个模型信息组中,查找确定出的商品类目对应的模型信息组;

从查找到的模型信息组的属性模型信息下的属性信息中,确定与所述查询关键字匹配的属性信息;

从确定出的属性信息中,提取出现次数达到设定门限值的属性信息;

将包含提取出的属性信息的商品节点列表作为搜索结果返回。

4. 如权利要求3所述的方法,其特征在于,在提取出对应属性信息出现次数达到设定门限值的属性信息之前,所述方法还包括:

如果在模型信息组中未查询出与所述查询关键字匹配的属性信息所对应的属性模型信息,则减少查询关键字的内容后,重新在查找到的模型信息组中查询与减少内容后的查询关键字匹配的属性信息所对应的属性模型信息。

5. 如权利要求4所述的方法,其特征在于,重新查询与减少内容后的查询关键字匹配的属性信息所对应的属性模型信息之后,所述方法还包括:

根据减少内容后的查询关键字更新查找到的模型信息组。

6. 如权利要求3所述的方法,其特征在于,服务器根据查询关键字确定的商品类目查找对应的模型信息组,具体包括:

服务器根据查询关键字中的属性信息所属的属性模型信息,从多个模型信息组中确定该属性模型信息所属的模型信息组。

7. 如权利要求3所述的方法,其特征在于,在模型信息组的属性模型信息中不存在与查询关键字匹配的属性信息时,所述方法还包括:

服务器从商品类目树中查询出与查询关键字匹配的所有信息,并返回包含该信息的所有节点的信息。

8. 如权利要求7所述的方法,其特征在于,所述方法还包括:

服务器收到对所述返回包含该信息的所有节点的信息中的商品信息的页面的访问信息时,生成日志文件,并根据该日志文件中的商品信息的属性信息和对应的所述查询关键字,更新模型信息组。

9. 一种记录用户访问商品信息的服务器,其特征在于,所述服务器包括:

日志生成模块,用于收到用户对包含商品信息的页面的访问信息时,生成日志文件,所述日志文件中包含访问页面中包含的商品信息的至少一个属性信息;

信息确定模块,用于根据生成的多个所述日志文件中包含的属性信息,分别确定同一商品类目对应的各个不同属性信息及每种属性信息出现的次数信息;

记录模块,用于将确定出的同一商品类目对应的各个不同属性信息以及每种属性信息出现的次数信息作为一模型信息组记录;其中:所述模型信息组包括属性模型信息,每一所述属性模型信息包括一属性信息及其对应出现的次数信息。

10. 一种利用权利要求 9 记录的用户访问商品信息对商品信息进行搜索的服务器,其特征在于,所述进行搜索的服务器包括:

类目确定模块,用于根据查询关键字,确定待查询商品信息所属的商品类目;

查找模块,用于在记录的各个模型信息组中,查找确定出的商品类目对应的模型信息组;

属性信息确定模块,用于从查找到的模型信息组的属性模型信息下的属性信息中,确定与所述查询关键字匹配的属性信息;

提取模块,用于从确定出的属性信息中,提取出现次数达到设定门限值的属性信息;

返回模块,用于将包含提取出的属性信息的商品节点列表作为搜索结果返回。

记录用户访问商品信息的方法及搜索方法和服务器

技术领域

[0001] 本申请涉及计算机技术领域,尤其涉及一种记录用户访问商品信息的方法和服务器,以及一种利用记录的用户访问商品信息的结果对商品信息进行搜索的方法和服务器。

背景技术

[0002] 用户在企业对个人用户 (Business To Customer, B2C) 网站或个人用户对个人用户 (Customer To Customer, C2C) 网站上进行商品搜索时,会通过网站提供的界面输入待查询的商品名称,网站根据用户输入的关键字进行搜索后向用户返回的结果主要有两种,一种是与用户待查询的商品相关的导航信息,另一种是用户待查询的商品的相关信息。

[0003] 目前,大量的商品信息按照商品类目名称以树的结构组织在一起,商品类目树保存在数据库相对应的数据表中,由人工对商品类目树中各节点的数据进行输入与维护, B2C 网站或是 C2C 网站中的每个商品信息的展示都属于商品类目树中某一个或多个节点。

[0004] 用户向 B2C 网站或是 C2C 网站进行商品信息查询时,如果网站向用户返回的结果是与用户待查询的商品相关的导航信息,则用户可以根据接收到的导航信息,沿商品类目树的路径自上而下定位至希望查询的搜索结果。如果网站向用户返回的结果是用户待查询的商品的相关信息,则网站将商品类目树中与用户待查询的商品相关的所有节点的信息返回给用户。

[0005] 由于目前的电子商务网站的商品数量非常庞大,有些网站的商品数量能够达到上亿规模,根据用户输入的待查询的商品名称搜索出的商品信息数量可能非常多,一方面,网站服务器向用户推送数量巨大的商品信息会占用服务器的大量系统资源以及网络带宽,另一方面,用户获得这些信息后,很难从网站返回的商品中准确、快速地定位出用户实际希望查询的商品。

[0006] 为了解决上述问题,目前的常规做法是限定向用户返回的商品类目数,通过减少向用户返回的商品信息,以减少向用户推送商品信息时对系统服务器的资源占用和网络的资源占用,并同时减少用户的查询时间。通过减少向用户返回的商品类目数的做法在一定程度上减少了资源的占用和用户的查询时间,但同时也可能将与用户查询相关度非常高的商品排除在外,导致向用户返回的查询结果不准确。

[0007] 综上所述,目前针对用户请求查询商品信息的搜索技术中,存在的对用户查询意图不明确,向用户返回的搜索结果中的信息与用户查询的相关度较低,导致用户的搜索结果较差的问题。

发明内容

[0008] 本申请的目的在于,提供一种记录用户访问商品信息的方法和服务器,用以解决现有技术中存在的对用户查询意图不明确的问题。

[0009] 一种记录用户访问商品信息的方法,所述方法包括:

[0010] 在用户每次访问包含商品信息的页面时生成日志文件,所述日志文件中包含访问

页面中包含的商品信息的至少一个属性信息；

[0011] 根据各个日志文件包含的属性信息，分别确定同一商品类目对应的各个不同属性信息、及每种属性信息出现的次数信息；以及

[0012] 将确定出的同一商品类目对应的各个不同属性信息、及每种属性信息出现的次数信息作为一模型信息组记录；

[0013] 其中：记录的该模型信息组中任一属性信息及其对应出现的次数信息作为该模型信息组中一个属性模型信息存在。

[0014] 一种记录用户访问商品信息的服务器，所述服务器包括：

[0015] 日志生成模块，用于收到用户对包含商品信息的页面的访问信息时，生成日志文件，所述日志文件中包含访问页面中包含的商品信息的至少一个属性信息；

[0016] 信息确定模块，用于根据生成的多个所述日志文件中包含的属性信息，分别确定同一商品类目对应的各个不同属性信息及每种属性信息出现的次数信息；

[0017] 记录模块，用于将确定出的同一商品类目对应的各个不同属性信息以及每种属性信息出现的次数信息作为一模型信息组记录；其中：所述模型信息组包括属性模型信息，每一所述属性模型信息包括一属性信息及其对应出现的次数信息。

[0018] 本申请记录用户访问包含商品信息的页面时生成日志文件，并对各个日志文件中的属性信息进行分析，将确定出的同一商品类目对应的各个不同属性信息、及每种属性信息出现的次数信息作为一模型信息组记录，构建用户对商品访问意图的模型，明确了设定时间内用户对商品的访问意图。

[0019] 本申请的另一目的在于，提供一种利用记录的用户访问商品信息对商品信息进行搜索的方法和服务器，用以解决现有技术中存在的向用户返回的搜索结果与用户查询的相关度较低，导致用户的搜索结果较差的问题。

[0020] 一种对商品信息进行搜索的方法，所述方法包括：

[0021] 根据用户输入的查询关键字，确定用户待查询商品信息所属的商品类目；

[0022] 在记录的各个模型信息组中，查找到确定出的商品类目对应的模型信息组；

[0023] 从查找到的模型信息组的属性模型信息下的属性信息中，确定与所述查询关键字匹配的属性信息；

[0024] 从确定出的属性信息中，提取出出现次数达到设定门限值的属性信息；

[0025] 将包含提取出的属性信息的商品节点列表返回给用户，其中商品节点列表中包含了与提取出的属性信息内容相同的至少一个包含商品信息的页面信息。

[0026] 一种对商品信息进行搜索的服务器，所述进行搜索的服务器包括：

[0027] 类目确定模块，用于根据查询关键字，确定待查询商品信息所属的商品类目；

[0028] 查找模块，用于在记录的各个模型信息组中，查找确定出的商品类目对应的模型信息组；

[0029] 属性信息确定模块，用于从查找到的模型信息组的属性模型信息下的属性信息中，确定与所述查询关键字匹配的属性信息；

[0030] 提取模块，用于从确定出的属性信息中，提取出现次数达到设定门限值的属性信息；

[0031] 返回模块，用于将包含提取出的属性信息的商品节点列表作为搜索结果返回。

[0032] 本申请利用已记录的用户对商品信息的访问意图对用户的搜索请求进行分类,查询出与用户的查询意图相关性较高的属性信息,并将包含查询出的属性信息的商品节点列表返回给用户,使用户得到高相关度的搜索结果,能够快速、准确的定位出希望获得的商品信息。在减少用户的查询时间和准确的定位出希望获得的商品信息的同时,有效减少了向用户推送商品信息时对系统服务器的资源占用和网络的资源占用。

附图说明

[0033] 图 1 为本申请实施例一确定商品的属性信息数量的方法示意图;

[0034] 图 2 为本申请实施例二进行商品搜索的方法示意图;

[0035] 图 3 为本申请实施例三记录用户访问商品信息的服务器结构示意图;

[0036] 图 4 为本申请实施例三对商品信息进行搜索的服务器结构示意图。

具体实施方式

[0037] 本申请提出对一段时间内用户对商品信息的访问进行记录,根据记录的汇总数据确定用户对商品信息的查询意图分布情况,得到用户对商品信息访问意图,进而利用得到的用户对商品信息的访问意图对用户的搜索请求进行分类,将与用户搜索的商品信息相关性较高的包含商品信息页面链接的商品信息节点列表返回给用户,让搜索结果更加接近用户的真实意图,使得搜索结果更准确。

[0038] 下面结合说明书附图对本申请的方案进行详细说明。

[0039] 实施例一

[0040] 本申请实施例一是记录用户访问商品信息的方法,如图 1 所示,包括以下步骤:

[0041] 步骤 101:服务器在用户每次访问包含商品信息的页面时生成日志文件。

[0042] 本实施例一中涉及的服务器是指能够在用户每次访问页面时,为本次访问事件生成日志文件的设备。该服务器可以是与提供商品信息页面的服务器集成在一起,也可以是独立于提供商品信息页面的服务器。

[0043] 本步骤的具体执行方式如下:

[0044] 预先在提供商品信息的页面中添加一条可以连接到生成日志文件的服务器(简称“日志服务器”)的链接,当用户通过搜索结果访问该页面或是通过其他方式(如通过商品类目树)访问该页面后,用户对该页面的每一次点击都由服务器生成一条日志文件,并将生成的该日志文件通过在页面中添加的链接保存到日志服务器中。本步骤中,涉及的日志服务器可以是独立的数据存储设备,也可以是存储商品类目树的数据库中专门用于存储日志文件的存储设备。

[0045] 本申请中涉及的日志文件中包含页面显示的商品信息的至少一个属性信息,所述属性信息包括商品品牌信息、商品型号信息、商品颜色信息或商品所属类目 ID 信息等。例如,用户访问一个提供的商品为手机的 Web 页面时,服务器针对该用户的这次访问生成的日志文件可以包含以下属性信息:手机的品牌信息为“ABC”,手机的型号信息为“123”,手机的颜色信息为“红色”,所属类目 ID 的信息为“手机”。

[0046] 如果用户是通过搜索引擎提供的导航信息沿商品类目树的路径自上而下到达的访问页面,则日志文件中还包含用户向搜索引擎提供的搜索关键字信息。例如,用户搜索的

关键字为“ABC 红色”，根据搜索引擎服务器返回的导航信息沿商品类目树访问提供品牌为“ABC”、型号为“123”、且颜色为“红色”的手机的 Web 页面时，则生成的日志文件可以包含以下属性信息：手机的品牌信息为“ABC”，手机的型号信息为“123”，手机的颜色信息为“红色”，搜索关键字信息为“ABC 红色”等。由于日志文件是在用户点击商品类目页面或是商品信息页面时由服务器自动生成的文本文件，因此，日志服务器中存储的各个日志文件中包含的用户输入的搜索关键字信息的格式可能会不统一，这种情况下，可以对存储的各个日志文件中的用户输入的关键字进行归一化处理，归一化处理的方式包括但不限于：去除不必要的词语、去除多余的空格、大小写字母的转换、全角半角的转换、繁体简体的转换、标点的转换和中文数字的转换等。

[0047] 步骤 102：根据设定时间长度内生成的各个日志文件包含的属性信息，分别确定同一商品类目对应的各个不同属性信息、及每种属性信息出现的次数信息。

[0048] 在本步骤中，可以对设定时间长度（如 24 小时）内生成的日志文件进行统计分析，确定用户在该设定时间长度内对商品信息的查询意图。在设定时间长度内生成的日志文件可以是多个用户访问提供商品信息的 Web 页面时生成的日志文件。

[0049] 在本步骤中，由于日志文件中包含的信息繁多，因此，可以通过支持向量机（support vector machine）对属性信息进行分类，产生分类数据。

[0050] 步骤 103：将确定出的同一商品类目对应的各个不同属性信息以及每种属性信息出现的次数信息作为一模型信息组记录。

[0051] 其中：记录的该模型信息组中任一属性信息及其对应出现的次数信息作为该模型信息组中一个属性模型信息存在。属性模型信息包括：商品品牌模型信息、商品型号模型信息和商品颜色模型信息等。例如，如果一条日志文件中包含的属性信息为商品品牌信息、商品型号信息、商品颜色信息和商品类目信息，则该商品类目信息对应的模型信息组中，商品品牌模型信息包括商品品牌信息及其出现次数，商品型号模型信息包括商品型号信息及其出现次数，商品颜色模型信息包括商品颜色信息及其出现次数。

[0052] 由于每一条日志文件是在用户访问提供某一商品信息的 Web 页面时生成的，因此，每一条日志文件对应一件商品信息。但是，某些商品信息中的一个或多个属性信息的内容可能是相同的，但这些商品信息表示的商品所属类目 ID 不同（如相同品牌的手机和电脑，其所属商品类目不同，但作为属性信息的商品品牌信息相同），因此，可以根据商品信息表示的商品所属类目来确定日志文件对应的类目。

[0053] 由于每一商品类目对应一个模型信息组，因此，将每个类目对应的模型信息组集合在一起，成为表示用户在设定时间长度内对商品信息访问意图的模型。

[0054] 下面以设定时间长度内生成 N 条日志文件为例，说明实施例一的具体实现方式：

[0055] 对生成的 N 条日志文件（1, 2, ..., n0, n1, n2, n3, ..., N）依次进行分析，确定每一条日志文件包含的属性信息，不断训练各商品类目对应的模型信息组，假设通过对第 1 条～第 n0 条日志文件的训练，得到的模型信息组如表 1 所示：

[0056]

模型信息组：类目	商品品牌模型信息 商品品牌信息：数量	商品型号模型信息 商品型号信息：数量	商品颜色模型信息 商品颜色信息：数量
模型信息组 1：手机	ABC: 2	123: 6	红色: 10
	DEF: 8	456: 20	黑色: 60
模型信息组 2：女裙	abc: 10	S: 30	白色: 45
	def: 6	M: 35	绿色: 50

[0057] 表 1

[0058] 假设：日志文件 n1 是访问提供某一款手机信息的 Web 页面时生成的日志文件，包含的属性信息为：“商品品牌信息：ABC”，“商品型号信息：123”，“商品颜色信息：红色”。

[0059] 日志文件 n2 是访问提供另一款手机信息的 Web 页面时生成的日志文件，包含的属性信息为：“商品品牌信息：DEF”，“商品型号信息：456”，“商品颜色信息：红色”。

[0060] 日志文件 n3 是访问提供一款女裙信息的 Web 页面时生成的日志文件，包含的属性信息为：“商品品牌信息：abc”，“商品型号信息：S”，“商品颜色信息：白色”。

[0061] 对上述第 n1 ~ 第 n3 条日志文件分析后，在表 1 的基础上进一步得到表 2 所示的模型信息组：

[0062]

模型信息组：类目	商品品牌模型信息 商品品牌信息：数量	商品型号模型信息 商品型号信息：数量	商品颜色模型信息 商品颜色信息：数量
模型信息组 1：手机	ABC: 3	123: 7	红色: 12
	DEF: 9	456: 21	黑色: 60
模型信息组 2：女裙	abc: 11	S: 31	白色: 46
	def: 6	M: 35	绿色: 50

[0063] 表 2

[0064] 类似地，在第 n3 条日志文件之后，可以继续利用第 n4 ~ 第 N 条日志消息不断更新表 2。表 2 所示的多个模型信息组的集合可以表示在设定时间长度内用户对多种类目商品的访问意图的模型。

[0065] 在对表 2 所示的用户对商品信息访问意图的模型进行存储时，不仅需要存储每一个模型信息组中的内容，还需要存储每一个模型信息组与类目的对应关系。

[0066] 在实施例一的方案中，对设定时间长度内保存在日志服务器中的日志信息，可以按照生成的时间先后顺序进行分析，训练得到模型信息组；也可以不分生成的先后顺序，对全部日志信息中的属性信息进行统一分析，训练得到模型信息组。

[0067] 通过实施例一的方案，对设定时间长度内用户对商品信息的访问，以日志文件的形式进行记录，并记录根据汇总数据确定的用户对商品信息的查询意图，以构建设定时间长度内用户对商品信息的访问意图的模型，从而确定设定时间长度内的用户访问意图。

[0068] 实施例二

[0069] 本申请实施例二利用实施例一记录的用户访问商品信息对商品信息进行搜索的方法,如图 2 所示,包括以下步骤:

[0070] 步骤 201:服务器根据接收的用户输入的查询关键字,确定用户待查询商品信息所属的商品类目。

[0071] 本实施例二中涉及的服务器是能够根据用户输入的关键字进行商品信息搜索的服务器,可以与实施例一中涉及的服务器集成在一起,也可以分别独立设置。

[0072] 步骤 202:在记录的各个模型信息组中,查找到确定出的商品类目对应的模型信息组。

[0073] 由于在实施例一的方案中,表 2 所示的用户对商品信息访问意图的模型中,每一个模型信息组与对应的商品类目保存在一起,因此,服务器在接收到用户输入的关键字时,可以根据所述关键字确定待查询的商品信息所属的商品类目,进而确定该商品类目 ID 对应的模型信息组。

[0074] 例如:用户向服务器输入的关键字为“手机、DEF”,则确定用户待查询商品信息商品类目是手机,且商品品牌信息是 DEF,对应表 2 模型中的模型信息组 1。在实施例二中,用户输入的查询关键字中也不限于包括待查询商品类目,还可以包括待查询商品的属性信息。

[0075] 步骤 203:从查找到的模型信息组的属性模型信息下的属性信息中,确定与所述查询关键字匹配的属性信息。

[0076] 假设用户输入的查询关键字是“手机、DEF”,则查找出的模型信息组为模型信息组 1,模型信息组 1 的属性模型信息包括商品品牌模型信息、商品型号模型信息和商品颜色模型信息,每一个属性模型信息又进一步包括属性信息。与查询关键字“DEF”匹配的属性信息包括:商品品牌信息“DEF”、商品型号信息“123、456”,商品颜色信息:“红色、黑色”。

[0077] 如果在本步骤中能够查询出与所述查询关键字匹配的属性信息,则表示用户输入的查询关键字是高频词,可以直接利用实施例一得到的模型信息组进行查询。

[0078] 如果在本步骤中未能够查询出与所述查询关键字匹配的属性信息,表示该用户输入的查询关键字是低频词,则可以减少查询关键字的内容后重新确定的属性信息中查询与减少内容后的查询关键字匹配的属性信息。例如:用户输入的查询关键字是“手机、ABD”,由于在实施例一中得到的模型信息组 1 中没有“ABD”品牌的手机,因此,减少查询关键字中的“ABD”,直接将“手机”作为查询关键字重新查询,以确定与更新后的查询关键字匹配的属性信息。

[0079] 再重新查询匹配的属性信息之后,确定“ABD”为低频词,则可以进一步利用该低频词进行实时对用户的查询意图分析,利用低频词更新表 2 中模型信息组 1 的内容,得到表 3 所示的用户对商品信息访问意图的模型:

[0080]

模型信息组: 类目	商品品牌模型信息 商品品牌信息: 数量	商品型号模型信息 商品型号信息: 数量	商品颜色模型信息 商品颜色信息: 数量
模型信息组 1: 手机	ABC: 3	123: 7	红色: 12
	DEF: 9	456: 21	黑色: 60
	ABD: 1		
模型信息组 2: 女裙	abc: 11	S: 31	白色: 46
	def: 6	M: 35	绿色: 50

[0081] 表 3

[0082] 步骤 204 :服务器从确定出的属性信息中,提取出出现次数达到设定门限值的属性信息。

[0083] 在本步骤中,为了向用户返回与查询相关度高的搜索结果,可以从步骤 203 中查询出的属性信息中进一步提取出在之前的设定时间长度内用户的查询意图较高的商品信息的属性信息,例如,在步骤 203 中查询出的属性信息包括:商品品牌信息“DEF”、商品型号信息“123、456”,商品颜色信息“红色、黑色”。通过表 2 可以确定商品品牌信息为“DEF”,因此,不论“DEF”的数量多少,都要将“DEF”作为提取出的属性信息;商品型号信息“123”的数量为 7,商品型号信息“456”的数量为 21,因此,可以将“456”作为提取出的属性信息;商品颜色信息为“红色”的数量为 12,商品颜色信息为“黑色”的数量为 60,因此,可以将“黑色”作为提取出的属性信息。此时,最终得到提取出的属性信息包括:商品品牌信息“DEF”、商品型号信息“456”和商品颜色信息“黑色”。

[0084] 步骤 205 :服务器将包含提取出的属性信息的商品节点列表返回给用户。

[0085] 商品节点列表中包含了与提取出的属性信息内容相同的至少一个包含商品信息的页面信息,如链接地址信息。

[0086] 由于商品信息按照商品信息类目树的形式保存在数据库中,因此,可以将包含商品品牌“DEF”、商品型号“456”和商品颜色“黑色”这些属性信息的部分或包含这些属性信息的商品节点列表按照 XML 的格式返回给用户。

[0087] 在用户通过商品节点列表中包括的商品页面的链接信息访问某一商品页面时,在该商品页面中还可进一步包含显示的商品的卖方信息、价格趋势信息和买方反馈信息中的一种或多种。

[0088] 在本实施例二的方案中,如果用户在步骤 201 中输入的查询关键字中不包含待查询商品所属的商品类目,则从表 2 的模型信息组中查询是否存在与查询关键字内容匹配的属性模型信息,如果存在,则可以根据查询出的属性模型信息确定该属性模型信息所属的模型信息组,进而确定出待查询商品所属的商品类目;否则,可以按照所述查询关键字查询出与用户待查询的商品的相关信息,并将查询出的相关信息返回给用户,并在用户访问所述相关信息指示的页面时生成日志文件,并利用生成的日志文件更新表 2 的模型信息组,其中:所述查询关键字的内容将作为更新后的模型信息组内的一部分内容。

[0089] 例如:如果用户在步骤 201 中输入的查询关键字为“ABC”,则从表 2 中模型信息组的内容可以确定待查询商品所属的商品类目为手机。

[0090] 如果用户在步骤 201 中输入的查询关键字为“ABD”，则从表 2 中无法确定待查询商品所属的商品类目，因此，服务器从保存在数据库中的商品类目树中查询出与“ABD”匹配的所有相关信息，并包含该相关信息的所有节点的信息返回给用户。用户获得服务器返回的信息后，在每次访问返回的信息指示的页面时，按照实施例一的方案生成日志文件。例如，用户访问一个提供的商品为 ABD 品牌的手机 Web 页面时，服务器针对该用户的这次访问生成的日志文件至少可以包含以下属性信息：手机的品牌信息为“ABD”，则可以根据当前生成的日志文件更新表 2，得到如表 3 所示的模型信息组：

[0091]

模型信息组：类目	商品品牌模型信息 商品品牌信息：数量	商品型号模型信息 商品型号信息：数量	商品颜色模型信息 商品颜色信息：数量
模型信息组 1：手机	ABC: 3	123: 7	红色: 12
	DEF: 9	456: 21	黑色: 60
	ABD: 1		
模型信息组 2：女裙	abc: 11	S: 31	白色: 46
	def: 6	M: 35	绿色: 50

[0092]

表 3

[0093] 在上述实例中，用户输入的查询关键字“ABD”可能是商品类目树中真实存在的商品属性信息的内容，也可能是用户在输入查询关键字时的误输入，如用户实际希望输入的查询关键字是“ABC”，但在输入时出现错误导致输入“ABD”，在按照上述实例中的方式向用户返回商品类目树中与用户待查询的商品相关的所有节点的信息后，用户访问页面时生成日志文件中应当包含该页面实际的属性信息以及用户输入的查询关键字。

[0094] 例如：用户访问的是提供的商品为 ABC 品牌的手机 Web 页面，此时，服务器针对该用户的这次访问生成的日志文件至少可以包含以下属性信息：手机的品牌信息为“ABC”和“ABD”，则可以根据当前生成的日志文件更新表 2，得到如表 4 所示的模型信息组：

[0095]

模型信息组：类目	商品品牌模型信息 商品品牌信息：数量	商品型号模型信息 商品型号信息：数量	商品颜色模型信息 商品颜色信息：数量
模型信息组 1：手机	ABC: 4	123: 7	红色: 12
	DEF: 9	456: 21	黑色: 60
	ABD: 1		
模型信息组 2：女裙	abc: 11	S: 31	白色: 46
	def: 6	M: 35	绿色: 50

[0096]

表 4

[0097] 在表 4 所示的模型信息组中，如果 ABD 是用户的误输入，则在利用表 4 执行本发明实施例二的方案时，由于误输入“ABD”对应的数量较少，达不到设定门限值，因此，在用户正

确输入查询关键字时,误输入不会影响查询结果的准确性;如果 ABD 不是用户的误输入,而是一种新式的手机品牌,则后续当有用户请求查询“ABD”时,可以按照表 4 为用户提供准确的查询结果。

[0098] 通过本实施例二的方案,利用已记录的用户对商品信息的访问意图对用户的搜索请求进行分类,查询出与用户的查询意图相关性较高的属性信息,使用户得到高相关度的搜索结果,从而能够快速、准确的定位出希望获得的商品信息。

[0099] 实施例三

[0100] 本申请实施例三还提供一种记录用户访问商品信息的服务器,如图 3 所示,所述服务器包括:日志生成模块 11、信息确定模块 12 和记录模块 13,其中:

[0101] 日志生成模块 11 用于收到用户对包含商品信息的页面的访问信息时,生成日志文件,所述日志文件中包含访问页面中包含的商品信息的至少一个属性信息;

[0102] 信息确定模块 12 用于根据生成的多个所述日志文件中包含的属性信息,分别确定同一商品类目对应的各个不同属性信息及每种属性信息出现的次数信息;

[0103] 记录模块 13 用于将确定出的同一商品类目对应的各个不同属性信息以及每种属性信息出现的次数信息作为一模型信息组记录;其中:所述模型信息组包括属性模型信息,每一所述属性模型信息包括一属性信息及其对应出现的次数信息。

[0104] 本实施例中涉及的商品信息的属性信息和属性模型信息与实施例一中定义相同。

[0105] 实施例四

[0106] 本申请实施例四在利用实施例三的服务器记录的用户访问商品信息基础上,提出一种对商品信息进行搜索的服务器,如图 4 所示,进行搜索的服务器包括:类目确定模块 21、查找模块 22、属性信息确定模块 23、提取模块 24 和返回模块 25,其中:

[0107] 类目确定模块 21 用于根据查询关键字,确定待查询商品信息所属的商品类目;

[0108] 查找模块 22 用于在记录的各个模型信息组中,查找确定出的商品类目对应的模型信息组;

[0109] 属性信息确定模块 23 用于从查找到的模型信息组的属性模型信息下的属性信息中,确定与所述查询关键字匹配的属性信息;

[0110] 提取模块 24 用于从确定出的属性信息中,提取出现次数达到设定门限值的属性信息;

[0111] 返回模块 25 用于将包含提取出的属性信息的商品节点列表作为搜索结果返回。

[0112] 所述查找模块 22 具体用于根据查询关键字中的属性信息所属的属性模型信息,从多个模型信息组中确定该属性模型信息所属的模型信息组。

[0113] 属性信息确定模块 23 还用于在模型信息组中未查询出与所述查询关键字匹配的属性信息所对应的属性模型信息时,减少查询关键字的内容后,重新在查找到的模型信息组中查询与减少内容后的查询关键字匹配的属性信息所对应的属性模型信息。

[0114] 所述进行搜索的服务器还包括更新模块 26,用于根据减少内容后的查询关键字更新查找到的模型信息组。

[0115] 返回模块 25 还用于在模型信息组的属性模型信息中不存在与查询关键字匹配的属性信息时,从商品类目树中查询出与查询关键字匹配的所有信息,并返回包含该信息的所有节点的信息。

[0116] 所述进行搜索的服务器还包括日志生成模块 27, 用于收到对所述返回包含该信息的所有节点的信息中的商品信息的页面的访问信息时, 生成日志文件; 更新模块 26 还用于根据该日志文件中的商品信息的属性信息和对应的所述查询关键字, 更新模型信息组。

[0117] 实施例三中的记录用户访问商品信息的服务器和实施例四中的对商品信息进行搜索的服务器可以是独立的网络设备, 也可以是集成在一起的网络设备。

[0118] 本领域内的技术人员应明白, 本申请的实施例可提供为方法、系统、或计算机程序产品。因此, 本申请可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且, 本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0119] 本申请是参照根据本申请实施例的方法、设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器, 使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

[0120] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中, 使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品, 该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

[0121] 这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上, 使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理, 从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

[0122] 尽管已描述了本申请的优选实施例, 但本领域内的技术人员一旦得知了基本创造性概念, 则可对这些实施例做出另外的变更和修改。所以, 所附权利要求意欲解释为包括优选实施例以及落入本申请范围的所有变更和修改。

[0123] 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。

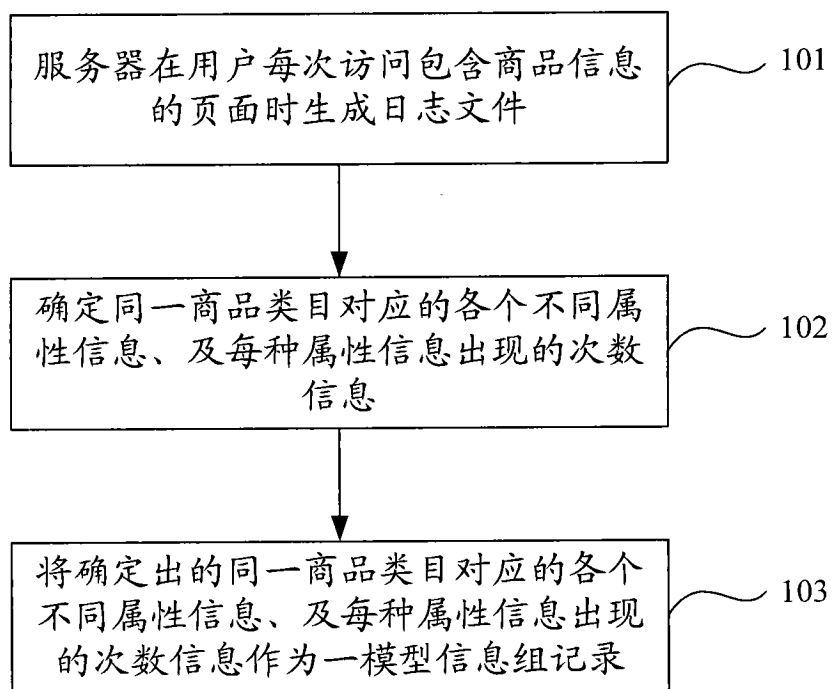


图 1

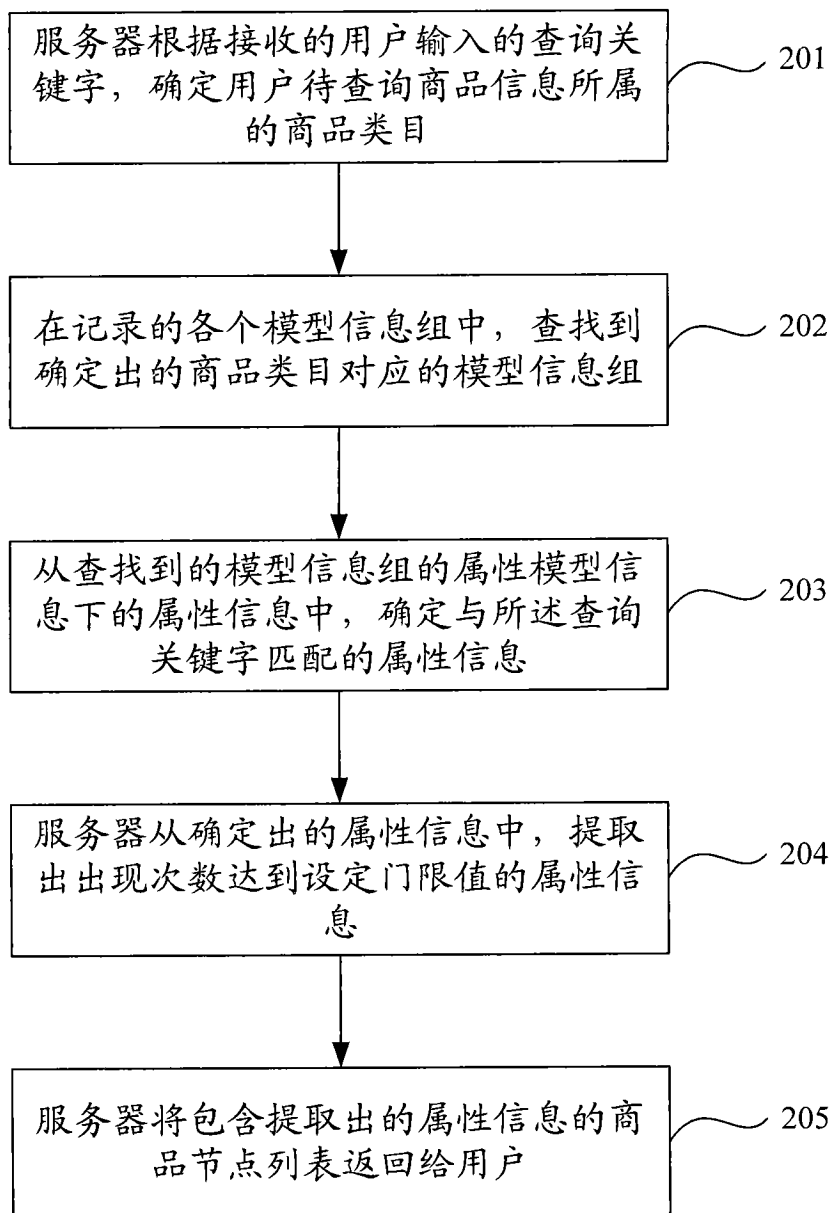


图 2

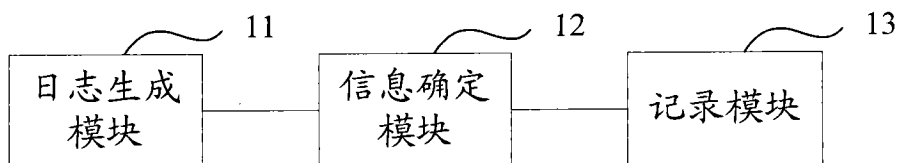


图 3

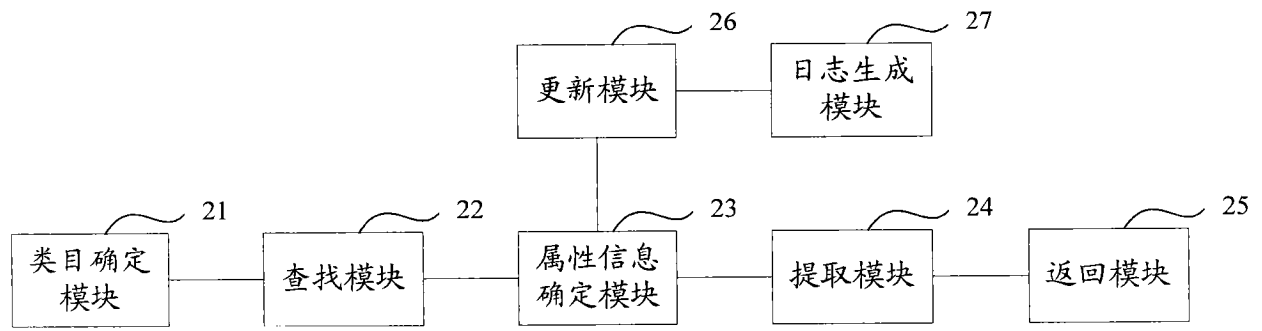


图 4