



(12) 发明专利申请

(10) 申请公布号 CN 102693223 A

(43) 申请公布日 2012. 09. 26

(21) 申请号 201110067079. 6

(22) 申请日 2011. 03. 21

(71) 申请人 潘燕辉

地址 100085 北京市海淀区学清路建清园小区 7 号楼 4 门 701

(72) 发明人 潘燕辉

(51) Int. Cl.

G06F 17/30 (2006. 01)

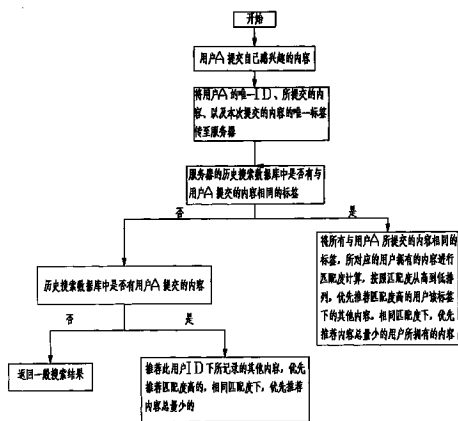
权利要求书 1 页 说明书 3 页 附图 2 页

(54) 发明名称

一种搜索方法

(57) 摘要

本发明公开了一种搜索方法,包括如下步骤,搜索用户在客户端提交自己感兴趣的内容,根据所述内容生成与所述内容对应的标记标签;将搜索用户的 ID、步骤 A1 所提交的内容以及所述标记标签上传至服务器,存储在历史搜索数据库中;获取所述的历史标签,及其对应的用户 ID,计算步骤 A1 所提交的内容与所述的历史标签下的内容的匹配度,按照匹配度进行排序,优先向所述搜索用户推荐匹配度较高的历史标签下的其他内容;相同匹配度下,优先推荐内容总量少的历史标签下的其他内容。



1. 一种搜索方法,其特征在于,包括如下步骤:

A1,搜索用户在客户端提交自己感兴趣的内容,根据所述内容生成与所述内容对应的标记标签;

A2,将搜索用户的 ID、步骤 A1 所提交的内容以及所述标记标签上传至服务器,存储在历史搜索数据库中;

A3,判断在所述历史搜索数据库中是否存在一历史标签,该历史标签下的内容与步骤 A1 所提交的内容相同或相近似;若为否,则执行步骤 A4;若为是,则执行步骤 A5;

A4,在所述历史搜索数据库中是否存在一历史内容,与步骤 A1 所提交的内容相同或相近似,若为是,则获取该历史内容所在的用户 ID 及其 ID 下所记录的其他内容,向所述搜索用户推荐该 ID 下所记录的其他内容;若为否,则向所述搜索用户推荐一般搜索结果;

A5,获取步骤 A3 所述的历史标签,及其对应的用户 ID,计算步骤 A1 所提交的内容与所述的历史标签下的内容的匹配度,按照匹配度进行排序,优先向所述搜索用户推荐匹配度较高的历史标签下的其他内容;相同匹配度下,优先推荐内容总量少的历史标签下的其他内容。

2. 根据权利要求 1 所述的搜索方法,其特征在于,所述步骤 A2 具体执行以下操作:

A21,服务器的历史搜索数据库中是否有搜索用户的 ID,若为是,则将步骤 A1 所提交的内容及其标记标签存储至历史搜索数据库中,若为否,则执行步骤 A22;

A22,询问搜索用户是否注册 ID,若为是,则进行 ID 注册为搜索用户建立 ID 后,则将步骤 A1 所提交的内容连同所述建立的 ID 及其标记标签存储至历史搜索数据库中;若为否,系统为搜索用户提供计算 ID,将步骤 A1 所提交的内容、所述计算 ID 及其标记标签存储至历史搜索数据库中。

3. 根据权利要求 1 所述搜索方法,其特征在于,所述客户端为计算机或智能手机终端或掌上电脑。

一种搜索方法

技术领域

[0001] 本发明涉及计算机技术领域,尤其涉及一种高命中率的信息搜索方法。

背景技术

[0002] 传统搜索引擎基本上有两种类型:

[0003] 1,根据所有用户针对某个选择的比例来进行推荐,当出现搜索结果时,选择多的结果优先推荐。当然其中还有其他参考的算法,比如结果的可信度,用户选择后的行为判断等。由于这种搜索基于大多数人的选择,所以命中率不高。

[0004] 2,情境搜索是综合考虑用户背景、兴趣爱好以及环境的智能化搜索,通过对用户意图的深入理解,在用户使用互联网服务的各种场景下提供给用户的最贴切的搜索服务。情境搜索包含 7 个要素 (6W&1H),它强调“以人 (Who) 为本”,也就是以用户为中心,根据其搜索行为的时间 (When)、地点 (Where)、输入 (What)、需求 (Want)、习惯 (How)、背景 (Why) 等因素,由情境计算得到最适合的搜索结果,再将这一结果通过用户的搜索情境直接呈现。但是这种搜索一定基于预先设计好的分类标准才能生效,比如用户关注阿凡达,搜索引擎必须知道阿凡达是部电影,而且是部科幻片,3D 大作,才有可能推荐同类的因素。

发明内容

[0005] 本发明所要解决的技术问题是针对现有技术的不足提供一种搜索方法——即同好搜索,所谓同好搜索,本质就是找到和搜索用户爱好相同或者相近的人,将他们喜欢的东西优先推荐给搜索用户,而不是传统上的分类推荐。

[0006] 本发明采用如下技术方案:

[0007] 一种搜索方法,包括如下步骤:

[0008] A1,搜索用户在客户端提交自己感兴趣的内容,根据所述内容生成与所述内容对应的标记标签;

[0009] A2,将搜索用户的 ID、步骤 A1 所提交的内容以及所述标记标签上传至服务器,存储在历史搜索数据库中;

[0010] A3,判断在所述历史搜索数据库中是否存在一历史标签,该历史标签下的内容与步骤 A1 所提交的内容相同或相近似;若为否,则执行步骤 A4;若为是,则执行步骤 A5;

[0011] A4,在所述历史搜索数据库中是否存在一历史内容,与步骤 A1 所提交的内容相同或相近似,若为是,则获取该历史内容所在的用户 ID 及其 ID 下所记录的其他内容,向所述搜索用户推荐该 ID 下所记录的其他内容;若为否,则向所述搜索用户推荐一般搜索结果;

[0012] A5,获取步骤 A3 所述的历史标签,及其对应的用户 ID,计算步骤 A1 所提交的内容与所述的历史标签下的内容的匹配度,按照匹配度进行排序,优先向所述搜索用户推荐匹配度较高的历史标签下的其他内容;相同匹配度下,优先推荐内容总量少的历史标签下的其他内容。

[0013] 所述的搜索方法,所述步骤 A2 具体执行以下操作:

[0014] A21,服务器的历史搜索数据库中是否有搜索用户的 ID,若为是,则将步骤 A1 所提交的内容及其标记标签存储至历史搜索数据库中,若为否,则执行步骤 A22;

[0015] A22,询问搜索用户是否注册 ID,若为是,则进行 ID 注册为搜索用户建立 ID 后,则将步骤 A1 所提交的内容连同所述建立的 ID 及其标记标签存储至历史搜索数据库中;若为否,系统为搜索用户提供计算 ID,将步骤 A1 所提交的内容、所述计算 ID 及其标记标签存储至历史搜索数据库中。

[0016] 所述搜索方法,所述客户端为计算机或智能手机终端或掌上电脑。

[0017] 这样利用“内容”、“标签”、“用户 ID”,搜索用户轻而易举就找到了与其具有相同或相近似的追求的用户,命中率高,可以应用于所有领域,而不必有事先的分类标准。它是一种先按照用户喜好去匹配拥有同样喜好的人,然后将拥有同样喜好的人的选择共享给拥有同样喜好的用户。而不必担心这种喜好是否是系统已知的分类。

附图说明

[0018] 图 1 为本发明搜索方法流程图;

[0019] 图 2 为本发明在服务器端建立历史搜索数据库的流程图。

具体实施方式

[0020] 以下结合具体实施例,对本发明进行详细说明。

[0021] 如图 1 所示为本发明同好搜索方法的流程图,包括如下步骤:

[0022] A1,搜索用户 A 在客户端提交自己感兴趣的内容,客户端根据其所提交的内容生成与该内容对应的一个标记标签,该标记标签可以任意生成,仅起标记作用;例如用户 A 提交的内容为书籍 B、书籍 F、书籍 T,参考表 1;

[0023] A2,将搜索用户 A 的 ID、步骤 A1 所提交的内容(书籍 B、书籍 F、书籍 T)以及所述标记标签上传至服务器,存储在历史搜索数据库中;

[0024] A3,判断在所述历史搜索数据库中是否存在这样的一个或多个历史标签,该历史标签下的内容与步骤 A1 所提交的内容相同或相近似;若为否,则执行步骤 A4;若为是,则执行步骤 A5;

[0025] A4,在所述历史搜索数据库中是否存在一历史内容,与步骤 A1 所提交的内容相同或相近似,若为是,则获取该历史内容所在的用户 ID 及其 ID 下所记录的其他内容,向所述搜索用户推荐该 ID 下所记录的其他内容;例如在表 1 中任何标签下都没有内容——书籍 S、书籍 G、书籍 K,但是在用户“ID-老潘”下存在一没有标签的项,该项下记录有书籍 S、书籍 G、书籍 K,则将该项下的其他书籍 L、书籍 H、书籍 D 推荐给用户。为否,则向所述搜索用户推荐一般搜索结果,即返回常规搜索引擎的搜索结果;

[0026] A5,获取步骤 A3 所述的历史标签,及其对应的用户 ID,计算步骤 A1 所提交的内容与所述的历史标签下的内容的匹配度,按照匹配度进行排序,优先向所述搜索用户推荐匹配度较高的历史标签下的其他内容;相同匹配度下,优先推荐内容总量少的历史标签下的其他内容,比如搜索用户提交了书籍 B、书籍 F、书籍 T,在表 1 中有两个标签——标签 ikjlojdsa 和标签 1238989 分别完全包含了书籍 B、书籍 F、书籍 T,则优先推荐这两个标签下的其他内容,例如书籍 J、书籍 L、书籍 Z,但是由于标签 1238989 下内容较少,所以优先推

荐书籍 Z, 其次推荐书籍 J 和书籍 L。而有一个标签——标签 mhiuwerf, 只包含了书籍 B、书籍 T, 则其匹配度较低, 作为次要推荐项。

[0027] 表 1 历史搜索数据库存储结构

[0028]

ID	历史标签	历史内容
ID-飞翔猴子	标签 asdhfkdsa	歌曲 A、歌曲 B、歌曲 C
	标签 lkhjdafff	商品 A、商品 B、商品 C
	标签 mhiuwerf	书籍 T、书籍 B、书籍 C、书籍 H、书籍 D、书籍 P、
ID-草草	标签 qertkdsau	书籍 A、书籍 B、书籍 C
	标签 ikjlojdsa	书籍 B、书籍 F、书籍 T、书籍 J、书籍 L
	标签 lkhoiuypoi	书籍 R、书籍 E、书籍 Q
ID-老潘	标签 noihjoi ju	电影 W、电影 Y、电影 O
	标签 1238989	书籍 B、书籍 F、书籍 T、书籍 Z、书籍 S、
	标签 3453289	商品 W、商品 B
		书籍 G、书籍 K、书籍 L、书籍 H、书籍 D

[0029] 本发明基于用户的输入, 先搜同样爱好的人, 然后再根据同样爱好的人拥有的选择或者信息来返回结果。在建立搜索数据库的时候, 普通搜索引擎都是利用一些程序在互联网上收集资料, 而同好搜索除了这种收集方式, 更加注意以用户为单位建立数据库。比如搜音乐, 将会在用户同意和知情的情况下, 将每个用户在服务器上建立一个唯一 ID(分为注册用户和非注册用户, 非注册用户将会根据用户电脑自动算出唯一码, 自动将用户进行分类), 同时, 在用户知情同意情况下, 将其所拥有的音乐信息上传至服务器存储在历史搜索数据库中。当用户使用同好搜索的时候, 也会将每次使用时所建立的标签和标签涉及到的音乐信息同样上传到此用户的数据库内。

[0030] 本发明的同好搜索应用范围极广, 前提是能搜集到能代表用户喜好的行为, 比如下载了某些资源, 比如经常访问的网页, 比如经常购买的物品, 比如收藏的某些东西, 比如经常进行交流的人群等等

[0031] 同样, 同好搜索和情景搜索可以有很好的融合, 比如当不能定位到具体项目时, 可以根据已有分类来进行大范围的同好搜索, 同样可以获得很好的结果。比如当用户经常访问某网店, 虽然没有购买具体的物品, 但是可以推荐给他同样经常光顾这个网店的人经常去的其他网店, 或者购买过的物品, 依然可能有很好的命中率

[0032] 应当理解的是, 对本领域普通技术人员来说, 可以根据上述说明加以改进或变换, 而所有这些改进和变换都应属于本发明所附权利要求的保护范围。

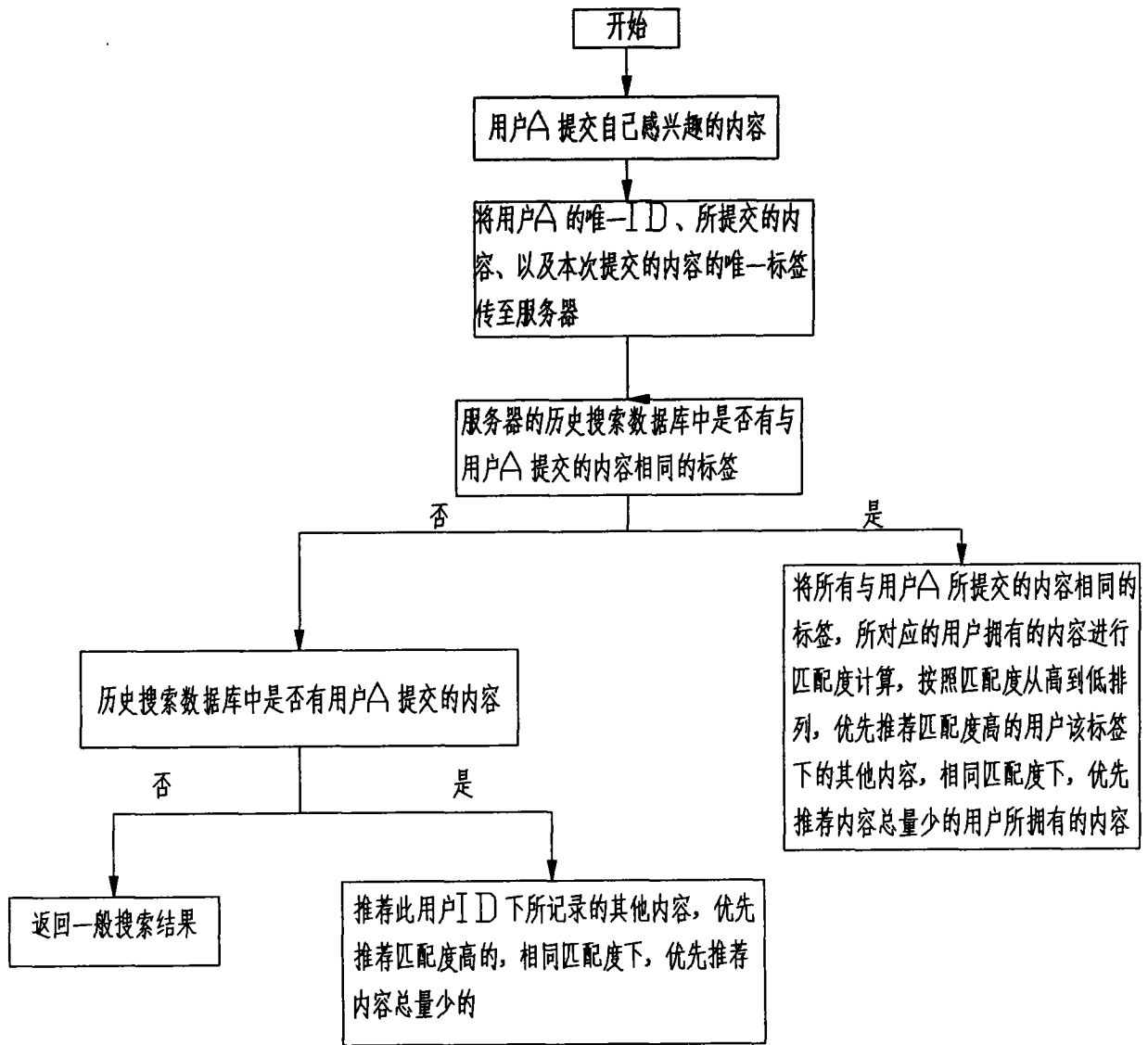


图 1

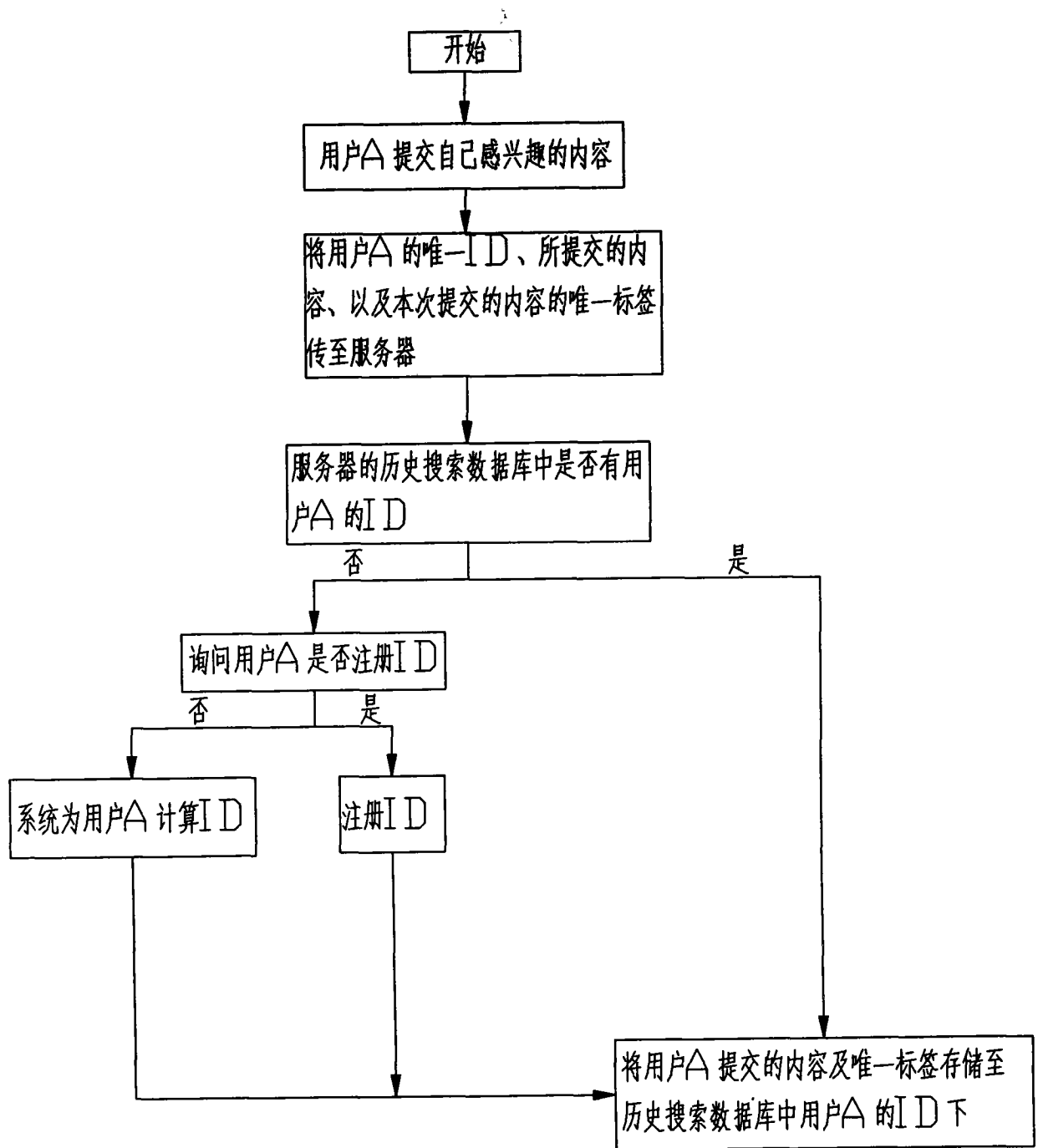


图 2