

(19) 中华人民共和国国家知识产权局



(12) 发明专利申请

(10) 申请公布号 CN 103324675 A

(43) 申请公布日 2013. 09. 25

(21) 申请号 201310197643. 5

(22) 申请日 2013. 05. 24

(71) 申请人 崔吉平

地址 230031 安徽省合肥市蜀山区红皖路
1088 号香山丽舍 27 栋 1304 室

申请人 李贵明

(72) 发明人 崔吉平 李贵明

(74) 专利代理机构 安徽合肥华信知识产权代理
有限公司 34112

代理人 余成俊

(51) Int. Cl.

G06F 17/30 (2006. 01)

权利要求书1页 说明书2页

(54) 发明名称

互联网个性化精确信息搜索及算法

(57) 摘要

本发明公开了一种互联网个性化精确信息搜索及算法，是结合互联网端信息关键字和用户端特征关键字，对互联网海量信息数据进行搜索筛选和最终排序显示的搜索算法。本发明除了具有目前通用搜索引擎利用搜索关键字(即信息特征)对互联网端信息来进行信息筛选和排序的功能，还将搜索结果与用户特征的关联度进行了详细的对比和数值表征，形成了更能符合互联网端信息关键字和用户端特征关键字的搜索结果，实现了个性化精确信息搜索的功能。

1. 互联网个性化精确信息搜索及算法,其特征在于,包括有以下步骤:

(1) 将互联网搜索引擎提供的搜索关键字分为互联网端信息关键字和用户端特征关键字两类,互联网端信息关键字即特征字符串,用户端特征关键字即用户身体特征、体型体态、健康状况、性别、性格、爱好、地域特征等各类要素;互联网端信息关键字由用户在搜索引擎页面输入,用户端特征关键字可通过用户在搜索引擎网站注册时填写个人档案,或在搜索引擎页面直接输入的方式,在进行搜索室与信息关键字同步提交至搜索引擎;

(2) 搜索结果的筛选算法既与搜索关键字的命中率、信息的点击量、更新程度、活跃程度等互联网端信息特征相关,也与用户端特征相关;

(3) 搜索结果的排序算法和规则是对搜索关键字、搜索结果所属网站特征,以及用户各特征要素进行数字化表征,根据餐饮、住宿、住房等各类不同的搜索类别,分别赋予互联网端信息特征如关键字、网站特征、点击量等及用户端特征以一定的搜索排序数值如权重、百分比等,符合普通搜索引擎对搜索关键字进行搜索的结果,需要对搜索结果进行用户端特征权重累加计算得出搜索结果匹配度指标,并将结果作为该搜索结果条目最终的排序依据;

(4) 个性化精确搜索算法得出的搜索结果,除显示搜索结果本身的内容以外,同时还显示根据用户端特征计算出的匹配指标数值,并给出与每一个用户端特征相关的特征匹配度详细数值列表,以便用户对搜索结果进行精确定位和筛选。

互联网个性化精确信息搜索及算法

技术领域

[0001] 本发明涉互联网信息搜索领域，具体涉及一种互联网端信息与用户端特征相结合的互联网个性化精确信息搜索及算法。

背景技术

[0002] 搜索引擎是指根据一定的策略、运用特定的计算机程序从互联网上搜集信息，在对信息进行组织和处理后，为用户提供检索服务，将用户检索相关的信息展示给用户的系统。搜索引擎包括全文索引、目录索引、元搜索引擎、垂直搜索引擎、集合式搜索引擎、门户搜索引擎与免费链接列表等。

[0003] 当用户以关键词查找信息时，搜索引擎会在数据库中进行搜寻，如果找到与用户要求内容相符的网站，目前各类搜索引擎采用的算法，通常根据网页中关键词的匹配程度、出现的位置、频次、链接质量等，计算出各网页的相关度及排名等级，然后根据关联度高低，按顺序将这些搜索结果及网页链接返回给用户。

[0004] 用户进行互联网信息搜索时，相对于互联网端信息特征，用户端特征也是对搜索结果进行筛选和排序的重要依据。

发明内容

[0005] 本发明是提供一种互联网个性化精确信息搜索及算法，是结合互联网端信息关键字和用户端特征关键字，对互联网海量信息数据进行搜索筛选和最终排序显示的搜索算法，从而提高用户搜索的命中率和准确性，实现个性化精确信息搜索功能。

[0006] 本发明采用的技术方案为：

互联网个性化精确信息搜索及算法，其特征在于，具体包括有以下步骤：

(1) 将互联网搜索引擎的搜索关键字分为互联网端信息关键字和用户端特征关键字两类，互联网端信息关键字即特征字符串，用户端特征关键字即用户身体特征、体型体态、健康状况、性别、性格、爱好、地域特征等各类要素；互联网端信息关键字由用户在搜索引擎页面输入，用户端特征关键字可通过用户在搜索引擎上填写用户档案的方式，或通过用户在搜索引擎页面直接输入的方式，在进行信息搜索时同步提交至搜索引擎；

(2) 将用户端特征关键字进行量化数据表征，即对各用户端特征关键字赋予一定的数值(如百分比或权重)，用户在选择不同的搜索类别(如餐饮、旅游、交通、住房、娱乐、医疗等)时，用户端特征关键字的数据表征具有不同的权重，与搜索类别具有强联系的用户端特征，其数据表征所占的权重较高，以突出搜索结果中与该类别相关的特征信息；

(3) 搜索引擎在收到搜索请求后，一方面将信息关键字与互联网海量信息进行关键字匹配；另一方面，将与信息关键字匹配的信息与用户端特征关键字进行比较，并根据用户端特征关键字数据表征及所占权重进行计算，得出该项用户的匹配度数值(如百分比或具体数值)；再将各项匹配度数值按照一定的权重系数累加后，形成该条搜索结果的用户端特征总匹配度数据；最后将该数据与传统搜索引擎排序规则结合，形成搜索结果的最终排序；

(4) 个性化精确搜索得到的搜索结果，在显示搜索结果基本内容的同时，一并显示该条目与各用户端的特征匹配数值，以直观的数据显示方式，更精确的表示搜索结果与用户端特征之间的关联度，从而利于用户对搜索条目进行进一步的筛选和比较。

[0007] 本发明的优点是：

本发明除了具有目前通用搜索引擎利用搜索关键字(即信息特征)对互联网端信息来进行信息筛选和排序的功能，还将搜索结果与用户特征的关联度进行了详细的对比和数值表征，形成了更能符合互联网端信息关键字和用户端特征关键字的搜索结果，实现了个性化精确信息搜索的功能。

具体实施方式

[0008] 互联网个性化精确信息搜索及算法，提供一种互联网络搜索引擎信息筛选、结果排序和显示方法，实现根据互联网用户不同个性特征与搜索关键字相结合的个性化精确信息搜索及算法；具体包括如下步骤：

(1) 利用自建的搜索引擎或租用其他搜索引擎的数据库，获得互联网各类信息基础数据，按照互联网用户各类特征属性，对数据库格式进行自定义和扩充，并填写相关信息，形成个性化精确搜索的基础数据库；

(2) 对用户端特征关键字进行数据表征，该步骤中，需要对用户搜索进行分类，用户端特征关键字的数据表征所占搜索结果权重，与搜索类别进行对应；例如，用户饮食偏好对餐饮类别搜索的影响权重，必然大于对住房类别搜索的权重；

(3) 搜索引擎在收到搜索请求后，一方面将信息关键字与互联网端海量信息进行关键字匹配；另一方面，将符合搜索关键字特征的信息与用户端特征关键字进行比较，并根据用户端特征关键字的数据表征(如百分比或具体数值)和所占权重，形成互联网端信息与用户端特征关键字的匹配度数据；最后，将各项匹配度数值进行累加，形成该条搜索结果的用户端特征匹配度数值；

(4) 将该用户端特征匹配度数据与传统搜索引擎排序规则结合，形成搜索结果的最终排序，在用户端界面进行显示。显示内容包括检索结果条目的基本信息，以及与用户端特征匹配度数据结果，供用户作为参考和依据；最终实现个性化精确搜索的功能。