

(12) 发明专利

(10) 授权公告号 CN 101355457 B

(45) 授权公告日 2011. 07. 06

(21) 申请号 200810115240. 0

(22) 申请日 2008. 06. 19

(73) 专利权人 腾讯科技(北京)有限公司
地址 100089 北京市海淀区海淀大街 38 号
银科大厦 16 层

(72) 发明人 邵荣防 张耀杰 杨娜

(74) 专利代理机构 北京德琦知识产权代理有限公司 11018
代理人 谢安昆 宋志强

11-12 行.

黄名选. 基于完全加权关联规则挖掘的查询扩展研究. 《中国优秀硕士学位论文全文数据库信息科技辑》. 2007, (第 05 期), 第 8 页倒数第 6-10 行, 第 9 页 13-18 行.

黄名选. 基于完全加权关联规则挖掘的查询扩展研究. 《中国优秀硕士学位论文全文数据库信息科技辑》. 2007, (第 05 期), 第 8 页倒数第 6-10 行, 第 9 页 13-18 行.

审查员 刘希

(51) Int. Cl.

H04L 12/26 (2006. 01)

H04L 12/24 (2006. 01)

G06F 17/30 (2006. 01)

(56) 对比文件

US 6718365 B1, 2004. 04. 06,

CN 101105815 A, 2008. 01. 16, 说明书第 11 页 10-20 行.

CN 1609859 A, 2005. 04. 27, 说明书第 4 页第

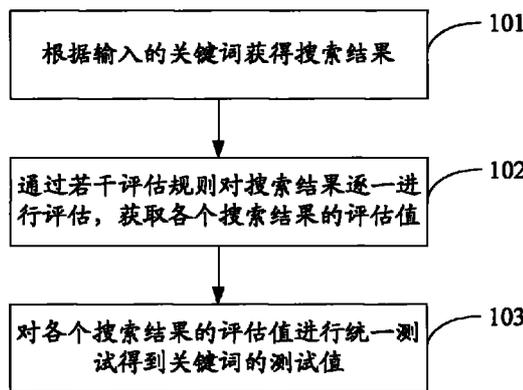
权利要求书 2 页 说明书 9 页 附图 4 页

(54) 发明名称

测试方法及测试设备

(57) 摘要

本发明公开了一种测试方法和测试设备, 所述方法包括: 根据输入的关键词获得搜索结果; 通过若干评估规则对所述搜索结果逐一进行评估, 获取各个搜索结果的评估值; 对所述各个搜索结果的评估值进行统一测试得到所述关键词的测试值。应用本发明对通过搜索引擎获取的测试用例进行测试时, 由于采用多种测试规则的组合对搜索结果进行自动的评测打分, 使得测试结果更加正确, 提高了测试结果的合理性和准确性, 通过该测试方式可以为后续搜索结果的排序优化提供了参考; 并且通过将上述测试规则通过软件实现, 可以对测试用例进行自动化测试, 降低了测试过程的复杂性, 提高了测试效率。



1. 一种测试方法,其特征在于,包括:
根据输入的关键词获得搜索结果;
通过若干评估规则对所述搜索结果逐一进行评估,获取各个搜索结果的评估值;
对所述各个搜索结果的评估值进行统一测试得到所述关键词的测试值。
2. 根据权利要求1所述的方法,其特征在于,还包括:构造关键词集合作为测试用例的输入数据。
3. 根据权利要求2所述的方法,其特征在于,所述构造关键词集合包括:根据词的类型构造关键词集合,或根据词的词性构造关键词集合。
4. 根据权利要求1所述的方法,其特征在于,所述通过若干评估规则对所述搜索结果逐一进行评估,获取各个搜索结果的评估值包括:
为所述若干评估规则分别设置权重值;
顺序提取一条搜索结果;
分别采用所述评估规则对所述搜索结果进行评估,得到对应每种评估规则的结果值;
根据所述权重值和所述结果值计算所述搜索结果的评估值。
5. 根据权利要求4所述的方法,其特征在于,所述评估规则包括:原文匹配规则、切词匹配规则、活跃度评估规则、特有因素评估规则、用户等级评估规则。
6. 根据权利要求1所述的方法,其特征在于,所述对各个搜索结果的评估值进行统一测试具体为:通过至少一种测试算法对所述各个搜索结果的评估值进行统一测试;
所述测试算法包括:最长不下降子序列算法、逆序对法。
7. 根据权利要求6所述的方法,其特征在于,当通过至少两种测试算法对各个搜索结果的评估值进行统一测试时,所述对各个搜索结果的评估值进行统一测试得到所述关键词的测试值包括:
为所述至少两种测试算法分别设置权重值;
分别采用所述测试算法对所述评估值进行测试,得到对应每种测试算法的结果值;
根据所述权重值和所述结果值计算所述关键词的测试值。
8. 一种测试设备,其特征在于,包括:
搜索结果获取单元,用于根据输入的关键词获得搜索结果;
评估单元,用于通过若干评估规则对所述搜索结果逐一进行评估,获取各个搜索结果的评估值;
测试单元,用于对所述各个搜索结果的评估值进行统一测试得到所述关键词的测试值。
9. 根据权利要求8所述的设备,其特征在于,还包括:
构造单元,用于构造关键词集合作为测试用例的输入数据。
10. 根据权利要求8所述的设备,其特征在于,所述评估单元包括:
第一设置单元,用于为所述若干评估规则分别设置权重值;
提取单元,用于顺序提取一条搜索结果;
第一获取单元,用于分别采用所述评估规则对所述搜索结果进行评估,得到对应每种评估规则的结果值;
第一计算单元,用于根据所述权重值和所述结果值计算所述搜索结果的评估值。

11. 根据权利要求 8 所述的设备,其特征在于,所述测试单元具体用于,通过至少一种测试算法对所述各个搜索结果的评估值进行统一测试,得到所述关键词的测试值。

12. 根据权利要求 11 所述的设备,其特征在于,所述测试单元包括:

第二设置单元,用于有至少两种测试算法时,为所述至少两种测试算法分别设置权重值;

第二获取单元,用于分别采用所述测试算法对所述评估值进行测试,得到对应每种测试算法的结果值;

第二计算单元,用于根据所述权重值和所述结果值计算所述关键词的测试值。

测试方法及测试设备

技术领域

[0001] 本发明涉及网络技术领域,特别涉及一种对网络中搜索结果进行评估的测试方法及测试设备。

背景技术

[0002] 搜索引擎是根据一定的策略,运用特定的计算机程序搜集互联网上的信息,在对信息进行组织和处理后,为用户提供检索服务的系统。搜索引擎类型多样,可以是针对所有互联网内容进行搜索的系统,也可以是针对特殊内容进行专项搜索的系统。搜索引擎的运营商根据搜索引擎的类型,建立专门的评估系统,用于对搜索结果进行评估测试,以使搜索引擎能够不断完善,更好的满足用户需求。

[0003] 对搜索结果的评估测试主要考虑两个方面的内容,分别是搜索结果的单一性和搜索结果的相关性。通常搜索结果会以文本的形式顺序排列,而在对上述搜索结果进行逐一评测时,现有技术通常只采用单一规则,这个单一规则就是判断搜索结果与输入的字符串的匹配程度。但是,通过字符串的匹配程度进行评测,只能测试出搜索结果与关键词在文本上是否一致,而无法测试搜索结果与用户需求是否匹配。由此可见,现有通过单一测试规则对搜索结果进行评测的方式,将导致测试结果不正确,降低了测试结果的合理性和准确性。

发明内容

[0004] 本发明的目的在于提供一种测试方法及测试设备,以解决现有技术中对搜索结果评估的规则单一,降低了测试准确性的问题。

[0005] 为解决上述技术问题,本发明提供如下技术方案:

[0006] 一种测试方法,包括:

[0007] 根据输入的关键词获得搜索结果;

[0008] 通过若干评估规则对所述搜索结果逐一进行评估,获取各个搜索结果的评估值;

[0009] 对所述各个搜索结果的评估值进行统一测试得到所述关键词的测试值。

[0010] 还包括:构造关键词集合作为测试用例的输入数据。

[0011] 所述构造关键词集合包括:根据词的类型关键词集合,或根据词的词性构造关键词集合。

[0012] 所述通过若干评估规则对所述搜索结果逐一进行评估,获取各个搜索结果的评估值包括:

[0013] 为所述若干评估规则分别设置权重值;

[0014] 顺序提取一条搜索结果;

[0015] 分别采用所述评估规则对所述搜索结果进行评估,得到对应每种评估规则的结果值;

[0016] 根据所述权重值和所述结果值计算所述搜索结果的评估值。

[0017] 所述评估规则包括:原文匹配规则、切词匹配规则、活跃度评估规则、特有因素评

估规则、用户等级评估规则。

[0018] 所述对各个搜索结果的评估值进行统一测试具体为：通过至少一种测试算法对所述各个搜索结果的评估值进行统一测试；

[0019] 所述测试算法包括：最长不下降子序列算法、逆序对法。

[0020] 当通过至少两种测试算法对各个搜索结果的评估值进行统一测试时，所述对各个搜索结果的评估值进行统一测试得到所述关键词的测试值包括：

[0021] 为所述至少两种测试算法分别设置权重值；

[0022] 分别采用所述测试算法对所述评估值进行测试，得到对应每种测试算法的结果值；

[0023] 根据所述权重值和所述结果值计算所述关键词的测试值。

[0024] 一种测试设备，包括：

[0025] 搜索结果获取单元，用于根据输入的关键词获得搜索结果；

[0026] 评估单元，用于通过若干评估规则对所述搜索结果逐一进行评估，获取各个搜索结果的评估值；

[0027] 测试单元，用于对所述各个搜索结果的评估值进行统一测试得到所述关键词的测试值。

[0028] 还包括：

[0029] 构造单元，用于构造关键词集合作为测试用例的输入数据。

[0030] 所述评估单元包括：

[0031] 第一设置单元，用于为所述若干评估规则分别设置权重值；

[0032] 提取单元，用于顺序提取一条搜索结果；

[0033] 第一获取单元，用于分别采用所述评估规则对所述搜索结果进行评估，得到对应每种评估规则的结果值；

[0034] 第一计算单元，用于根据所述权重值和所述结果值计算所述搜索结果的评估值。

[0035] 所述测试单元具体用于，通过至少一种测试算法对所述各个搜索结果的评估值进行统一测试，得到所述关键词的测试值。

[0036] 所述测试单元包括：

[0037] 第二设置单元，用于有至少两种测试算法时，为所述至少两种测试算法分别设置权重值；

[0038] 第二获取单元，用于分别采用所述测试算法对所述评估值进行测试，得到对应每种测试算法的结果值；

[0039] 第二计算单元，用于根据所述权重值和所述结果值计算所述关键词的测试值。

[0040] 由以上本发明提供的技术方案可见，本发明根据输入的关键词获得搜索结果，通过若干评估规则对所述搜索结果逐一进行评估，获取各个搜索结果的评估值，对所述各个搜索结果的评估值进行统一测试得到所述关键词的测试值。应用本发明对通过搜索引擎获取的测试用例进行测试时，由于采用多种测试规则的组合对搜索结果进行自动的评测打分，使得测试结果更加正确，提高了测试结果的合理性和准确性，通过该测试方式可以为后续搜索结果的排序优化提供了参考；并且通过将上述测试规则通过软件实现，可以对测试用例进行自动化测试，降低了测试过程的复杂性，提高了测试效率。

附图说明

- [0041] 图 1 为本发明测试方法的一个实施例流程图；
[0042] 图 2 为本发明测试方法的另一个实施例流程图；
[0043] 图 3 为本发明测试方法的又一个实施例流程图；
[0044] 图 4 为本发明测试设备的一个实施例框图；
[0045] 图 5 为本发明测试设备的另一个实施例框图。

具体实施方式

[0046] 本发明的核心是提供了一种测试方法和测试设备,根据输入的关键词获得搜索结果,通过若干评估规则对搜索结果逐一进行评估,获取各个搜索结果的评估值,对各个搜索结果的评估值进行统一测试得到关键词的测试值,由于对搜索结果进行评估时综合了多种评估规则,相应提高了测试结果的准确性。

[0047] 为了使本技术领域的人员更好地理解本发明方案,并使本发明的上述目的、特征和优点能够更加明显易懂,下面结合附图和具体实施方式对本发明作进一步详细的说明。

[0048] 本发明测试方法的一个实施例流程如图 1 所示:

[0049] 步骤 101:根据输入的关键词获得搜索结果。

[0050] 步骤 102:通过若干评估规则对搜索结果逐一进行评估,获取各个搜索结果的评估值。

[0051] 步骤 103:对各个搜索结果的评估值进行统一测试得到关键词的测试值。

[0052] 本发明测试方法的另一个实施例流程如图 2 所示,该实施例示出了从构造测试用例开始到获取测试结果值的详细过程:

[0053] 步骤 201:构造关键词集合作为测试用例的输入数据。

[0054] 其中,可以根据词的类型构造关键词集合,例如,词的类型为体育,则构造的关键词集合中可以包括足球、篮球、羽毛球、乒乓球等;也可以根据词的词性构造关键词集合,例如,词的词性为动词,则构造的关键词集合中可以包括追逐、帮助、哭泣等。

[0055] 步骤 202:为不同的评估规则分别设置权重值,并为不同的测试算法分别设置权重值。

[0056] 其中,评估规则可以包括原文匹配规则、切词匹配规则、活跃度评估规则、特有因素评估规则、用户等级评估规则;测试算法可以包括最长不下降子序列算法、逆序对法。

[0057] 在设置权重值时,主要根据评估规则或测试算法的重要程度进行设置,也可以根据需要进行设置,即如果评估规则或测试算法的重要性高则为其分配的权重值较大,反之则较小。

[0058] 步骤 203:根据输入的关键词获得搜索结果。

[0059] 通常在搜索引擎中输入一个关键词后,会输出数量庞大的搜索结果,对于测试来说,通常只需要选择其中的前五十条进行评测即可。

[0060] 步骤 204:顺序提取一条搜索结果。

[0061] 步骤 205:分别采用不同的评估规则对该条搜索结果进行评估,得到对应每种评估规则的结果值。

[0062] 步骤 206 :根据分配的权重值和结果值计算该条搜索结果的评估值。

[0063] 具体的,可以将为每条评估规则分配的权重值与该条评估规则的结果值相乘,并将所有乘积累加即可获得该条搜索结果的评估值。

[0064] 步骤 207 :判断是否为输入关键词的最后一条搜索结果,若是,则执行步骤 208 ;否则,返回步骤 204。

[0065] 最后一条搜索结果并不是所有搜索结果的最后一条,如果设置需要评估的搜索结果为前五十条,则相应的最后一条搜索结果为第五十条搜索结果。

[0066] 步骤 208 :分别采用不同的测试算法对所有搜索结果的评估值进行测试,得到对应每种测试算法的结果值。

[0067] 步骤 209 :根据设置的权重值和结果值计算输入关键词的测试值。

[0068] 具体的,可以将为每种测试算法分配的权重值与该测试算法的结果值相乘,并将所有乘积累加即可获得该关键词的测试值。

[0069] 本发明测试方法的另一个实施例流程如图 3 所示,该流程特别结合搜索引擎日志中的关键词,对测试过程的具体应用进行详细描述 :

[0070] 步骤 301 :按照搜索引擎日志中词的类型构造测试用例关键词的集合。

[0071] 通过分析搜索引擎日志中关键词的主要类型,提取类型集合并对每种类型包含的关键词进行扩充,构造测试用例集合作为测试的输入数据。

[0072] 以网络中的日志为例,可以构造如下表 1 所示的测试用例,每个测试用例中包含词的类型及该类型对应的关键词 :

[0073] 表 1

[0074]

测试用例序号	类型	关键词
1	日志 / 心情	毕业心情、追忆似水年华、军训生活
2	音乐 / 影视	宝贝计划、变形金刚、大长今
3	汽车 / 手机	奔驰、诺基亚、摩托罗拉

[0075] 为每个测试用例对应的类型构造的关键词不局限于上表所列的数量,可以根据需要构造更多的关键词集合。

[0076] 步骤 302 :顺序对测试用例中的关键词进行搜索得到搜索结果。

[0077] 假设以上表 1 中测试用例 1 中的“毕业心情”作为关键词进行搜索,获得两条搜索结果。通常搜索结果以文本方式顺序排列,包含标题(关键词)和部分正文摘要,本实施例为了描述方便,以列表方式给出两条搜索结果,由于该实施例示出的是对网络中日志进行的搜索,因此每条搜索结果中除了包含标题和正文摘要外,还包含日志的浏览次数、回复次数、是否置顶、是否转载等,如下表 2 所示 :

[0078] 表 2

[0079]

标题	浏览	回复	置顶	转载	图片	等级	正文
毕业心情	0	68	是	否	有	高	毕业,是个沉重的动词毕业,是个让人一生难忘的名词毕业,是感动流泪的形容词每个人都有离别时的感受,在那一刻我们回忆起与同学每时每刻相处的日子.....
毕业的心情	0	0	否	否	无	低	马上面临毕业,真的很舍不得。刚来学校时,陌生的面孔,现在成为了朋友。几年来的相处使人有点难忘。毕业了,朋友难聚了。将来回忆起来,一定很美好.....

[0080] 步骤 303 :提取一条搜索结果。

[0081] 步骤 304 :采用原文匹配规则对搜索结果进行评估。

[0082] 原文匹配规则就是基于字符串匹配的方式,对文本进行关键词的匹配。文本分为标题文本和摘要文本,对于第一条搜索结果,标题文本中含有完整的“毕业心情”,其与关键词完全匹配,正文中不含有完整的“毕业心情”,其与关键词不完全匹配,这里不考虑关键词匹配的个数,只考虑是否能够匹配。通常标题文本匹配的重要性大于摘要文本数,因此可以为标题文本和摘要文本设置不同的匹配权重,根据设置的权重进行评估的公式可以为: $W_{title} \times B_{title} + W_{abs} \times B_{abs}$,其中, W_{title} 为标题文本的权重, W_{abs} 为摘要文本的权重, B_{title} 表示标题文本是否匹配, B_{abs} 标识摘要文本是否匹配,如果匹配则对应的值为 1,如果否则对应的值为 0,假设 W_{title} 与 W_{abs} 的比为 2 : 1,该权重可以根据测试需要进行灵活调整,按照上述公式计算得到的评估结果 $S(result1) = 2 \times 1 + 1 \times 0 = 2$ 。

[0083] 步骤 305 :采用切词匹配规则对搜索结果进行评估。

[0084] 切词匹配规则指将关键词分成不同的子关键词,将搜索到的文本分成若干词,用上述子关键词与由文本分成的若干词分别进行匹配,根据子关键词的匹配结果对搜索结果进行评估。

[0085] 关键词“毕业心情”分为子关键词“毕业”和“心情”,第一条搜索结果的文本分成若干词后的结果如下所示,为了示例清楚,不同的词之间用“|”分隔,其中标题文本分为“毕业|心情”,摘要文本分为“毕业|,是|个|沉重|的|动词|毕业|,是|个|让|人|一生|难忘|的|名词|毕业|,是|感动|流泪|的|形容词|每个|人|都|有|离别时|的|感受|,在|那|一刻|我们|回忆|起|与|同学|每时每刻|相处|的|日子.....”。为每个子关键词分别设置对应的权重,根据设置的权重进行评估的公式可以为: $W_{title} \times (B_{TxdDivResult[i]} \times \text{len}(TxdDivResult[i]) \div \text{len}(TxdResult)) + W_{abs} \times (B_{TxdDivResult[i]} \times \text{len}(TxdDivResult[i]) \div \text{len}(TxdResult))$,其中, W_{title} 为标题文本的权重, W_{abs} 为摘要文本的权重, $B_{TxdDivResult[i]}$ 表示文本是否与子关键词匹配,如果匹配则值为“1”,否则值为“0”, $\text{len}(TxdDivResult[i])$ 表示子关键词的长度, $\text{len}(TxdResult)$ 表示所有关键词的长度。

[0086] 对于不同的子关键词, 搜索结果的标题文本中切分后的词与子关键词“毕业”和“心情”均匹配, 摘要文本中切分后的词与子关键词“毕业”匹配, 而与“心情”不匹配。按照上述公式计算得到的评估结果 $S(\text{result}2) = 2 \times (1 \times 2 \div 4 + 1 \times 2 \div 4) + 1 \times (1 \times 2 \div 4 + 0 \times 2 \div 4) = 2.5$ 。

[0087] 步骤 306 : 采用活跃度评估规则对搜索结果进行评估。

[0088] 活跃度评估规则是指根据网络中该搜索结果对应的 URL (Uniform Resource Locator, 统一资源定位符) 被点击次数和回复次数的多少得到该搜索结果的评估结果。

[0089] 根据表 1 中所示, 第一条搜索结果被浏览的次数为 0, 被回复的次数为 68, 由于回复表示浏览该 URL 对应日志的用户对该日志的兴趣大于仅仅浏览该日志的用户, 因此为浏览设置的权重小于为回复设置的权重, 假设浏览和回复的权重比为 1 : 10, 则第一条搜索结果的评估值为 : $S(\text{result}3) = 1 \times 0 + 10 \times 68 = 680$ 。

[0090] 步骤 307 : 采用特有因素评估规则对搜索结果进行评估。

[0091] 由于本实施例是对网络中搜索日志的结果进行的评测, 对于日志来说, 有其特有的功能, 如表 1 中所示该搜索结果对应的日志是否被置顶、是否被转载或是否添加了图片都反应了用户对该日志的喜爱程度, 也就是说, 一条搜索结果对应的日志如果被转载和置顶则在评测时为其分配的权重应大于没有被转载和置顶的日志, 但是, 如果同一篇日志被转载的次数增多, 则该日志重复的次数也相应变大, 因此应加大其被过滤的可能性, 降低其权重分配; 而如果一条搜索结果对应的日志中加入了图片, 则该日志为美文的可能性就越大, 因此应该为具有图片的日志分配一定的权重。

[0092] 如表 1 所示, 第一条搜索结果对应的日志被置顶, 未被转载且具有图片, 假设为上述三种功能分配的权重比为 5 : -2 : 2, 其中为了减少由于转载而重复的日志排到搜索结果的前面, 将转载的权重设为负数。对于是否置顶、是否转载和是否具有图片来说, 如果其为是, 则对应的值为“1”, 如果其否, 则对应的值为“0”。根据设置的权重进行评估的公式可以为 : $W_{\text{top}} \times B_{\text{top}} + W_{\text{zz}} \times B_{\text{zz}} + W_{\text{photo}} \times B_{\text{photo}}$, 其中, W_{top} 表示为置顶分配的权重, B_{top} 表示是否置顶的值, W_{zz} 表示为转载分配的权重, B_{zz} 表示是否转载的值, W_{photo} 表示为图片分配的权重, B_{photo} 表示是否具有图片的值。按照上述公式计算得到的第一条搜索结果的评估结果 $S(\text{result}4) = 5 \times 1 - 2 \times 0 + 2 \times 1 = 7$ 。

[0093] 步骤 308 : 采用用户等级评估规则对搜索结果进行评估。

[0094] 由于网络中日志对应的用户因为其使用时间的长短或其它因素等具有不同的等级, 而对于等级较高的用户, 其所写日志在搜索结果中的排序应该在等级较低的用户前面, 因此为等级较高用户分配的权重应该高于为等级较低用户分配的权重。

[0095] 如表 1 所示, 对所有用户等级分为高和低两种, 其中为等级“高”分配的权重值为 1, 对等级“低”分配的权重值为 0, 对于第一条搜索结果来说, 其对应用户的等级为“高”, 相应的权重值就为“1”。因此根据设置的权重进行评估的公式可以为 : $S(\text{result}5) = W_{\text{user}}$, W_{user} 就表示相应等级的权重值, 按照上述公式计算得到的第一条搜索结果的评估结果 $S = 1$ 。

[0096] 步骤 309 : 按照预先为每种评估规则设置的权重获取第一条搜索结果的最终测试结果。

[0097] 将按照上述步骤 304 至步骤 308 的评估规则对第一条搜索结果进行评测得到的评

估结果值换算成百分制得分。

[0098] 采用原文匹配规则得到的评估结果的满分为“3”，而第一条搜索结果得到的评估结果值为“2”，相应换算为百分制后，第一条搜索结果的分值 $S_1 = 66.7$ ；采用切词匹配规则得到的评估结果的满分为“3”，而第一条搜索结果得到的评估结果值为“2.5”，相应换算为百分制后，第一条搜索结果的分值 $S_2 = 83.3$ ；对于活跃度评估规则，假设浏览 1000 次以上与 1000 次效果相同，回复 100 次以上与 100 次效果相同，即满分为 2000 分，第一条搜索结果得到的评估结果值为“680”，相应换算为百分制后，第一条搜索结果的分值 $S_3 = 34$ ；采用特有因素评估规则得到的评估结果的满分为“7”，而第一条搜索结果得到的评估结果值为“7”，相应换算为百分制后，第一条搜索结果的分值 $S_4 = 100$ ；采用用户等级评估规则得到的评估结果的满分为“1”，而第一条搜索结果得到的评估结果值为“1”，相应换算为百分制后，第一条搜索结果的分值 $S_5 = 100$ 。

[0099] 进一步，根据上述五个评估规则的重要程度，灵活地为其分别分配相应的权重，此处假设分配的权重比为 $W_1 : W_2 : W_3 : W_4 : W_5 = 1 : 4 : 2 : 2 : 1$ ，根据设置的权重进行评估的总公式可以为： $S = W_1 \times S_1 + W_2 \times S_2 + W_3 \times S_3 + W_4 \times S_4 + W_5 \times S_5 = 1 \times 66.7 + 4 \times 83.3 + 2 \times 34 + 2 \times 100 + 1 \times 100 = 768$ 。

[0100] 需要说明的是，上述步骤 304 至 308 的顺序可以变换，只要用所有的评估规则对搜索结果进行评估即可。另外，根据评测的需要，也可以仅选择其中的几条规则进行评测，由于不是按照单一的评估规则进行评估，因此对每一条搜索结果得到的评估值更客观和合理。

[0101] 步骤 310：判断是否测试完所有的搜索结果，若是，则执行步骤 311；否则返回步骤 303。

[0102] 步骤 311：通过所有搜索结果的评估结果值获取该关键词的测试结果。

[0103] 本实施例中为了示例方便，对关键词“毕业心情”仅示出了两条搜索结果，而在实际搜索过程中，对于同一输入关键词通常会输出成百上千条搜索结果，而这些搜索结果中往往前五十条搜索结果与关键词的匹配程度越高，也是用户所最为关注的。在获取该关键词的测试结果时，通常对获取的前五十条搜索结果分别按照步骤 303 至步骤 307 中的评估规则进行打分。

[0104] 为了描述获取关键词的测试结果的过程，可以分别采用最长不下降子序列法和逆序对法对某个关键词的搜索结果进行整体评测。下面假设对某个关键词共获得了 11 条搜索结果，这 11 条搜索结果的测试得分如下表 3 所示（以十分制打分为例）：

[0105] 表 3

[0106]

搜索结果	1	2	3	4	5	6	7	8	9	10	11
------	---	---	---	---	---	---	---	---	---	----	----

[0107]

结果打分	3	2	0	3	3	2	1	2	1	1	0
------	---	---	---	---	---	---	---	---	---	---	---

[0108] 最长不下降子序列法指在一个区间的数列中，按照降序排列的数的长度，该区间用 $f(a, b)$ 表示，例如，对于搜索结果 1 至 5 的区间 $f(1, 5)$ ，即 3、2、0、3、3 中，抽取按照降序排列的数为“3、3、3”，相应的降序排列的数的长度为 3。采用最长不下降子序列法对某个关

关键词进行评测时,可以忽略搜索结果中个别位置不正确的搜索结果,而是考虑某个区间数列中的整体排序结果。对于如表 3 中所示的关键词,分别获取搜索结果 1 至 5(用户最关注的搜索结果),即 $f(1,5)$ 的最长不上升子序列的长度为 3,搜索结果 1 至 10(第一页的搜索结果),即 $f(1,10)$ 的最长不上升子序列的长度为 7,搜索结果 1 至最后一个(通常为 50,表 3 中为 11),即 $f(1,11)$ 的最长不上升子序列的长度为 8,为上述不同区间分配对应的权重,假设上述三个区间 $f(1,5)$ 、 $f(1,10)$ 和 $f(1,11)$ 的权重比为 $w_1 : w_2 : w_3 = 10 : 5 : 1$,采用公式 $S(A) = w_1 \times f(1,5) + w_2 \times f(1,10) + w_3 \times f(1, \text{end})$ 计算该关键词的评估结果值为: $S(A) = 10 \times 3 + 5 \times 7 + 1 \times 8 = 73$ 。

[0109] 逆序对法是指在一个区间数列中,排列在前的数的序号小于排列在后的数的序号,而如果排列在前的数的值大于排列在后的数的值,则这两个数称为一个逆序对。例如对于搜索结果 1 至 5 的区间 $f(1,5)$,即 3、2、0、3、3 中,排列在前的数的值大于排列在后的数的值共有 3 对,分别为 3 和 2,3 和 0,2 和 0。采用逆序对法对某个关键词进行评测时,可以忽略搜索结果中给定区间内位置不合理的单个搜索结果值。对于如表 3 中所示的关键词,同样分别获取搜索结果 1 至 5(用户最关注的搜索结果),即 $f(1,5)$ 的逆序对的个数为 3,搜索结果 1 至 10(第一页的搜索结果),即 $f(1,10)$ 的逆序对的个数为 26,搜索结果 1 至最后一个(通常为 50,表 3 中为 11),即 $f(1,11)$ 的最长不上升子序列的长度为 35,为上述不同区间分配对应的权重,假设也设置上述三个区间 $f(1,5)$ 、 $f(1,10)$ 和 $f(1,11)$ 的权重比 $w_1' : w_2' : w_3' = 10 : 5 : 1$,采用公式 $S(B) = w_1' \times f(1,5) + w_2' \times f(1,10) + w_3' \times f(1, \text{end})$ 计算该关键词的评估结果值为: $S(B) = 10 \times 3 + 5 \times 26 + 1 \times 35 = 195$ 。

[0110] 进一步,由于采用上述两种方法获得了某关键词的不同评估结果值,可以按照两种方法的重要程度,为上述两种方法分配不同的权重值,这里假设两种方法的权重比 $W(A) : W(B) = 10 : 1$,则采用公式 $S = W(A) \times S(A) + W(B) \times S(B)$ 获得该关键词的最终评测结果值 $S = 10 \times 73 + 1 \times 195 = 925$ 。

[0111] 步骤 312:判断是否将测试用例中的所有关键词测试完毕,若是,则结束流程;否则返回步骤 302。

[0112] 与本发明测试方法的实施例相对应,本发明还提供了测试设备的实施例。

[0113] 本发明测试设备的一个实施例框图如图 4 所示,该测试设备包括:搜索结果获取单元 410、评估单元 420 和测试单元 430。

[0114] 其中,搜索结果获取单元 410 用于根据输入的关键词获得搜索结果;评估单元 420 用于通过若干评估规则对所述搜索结果逐一进行评估,获取各个搜索结果的评估值;测试单元 430 用于对所述各个搜索结果的评估值进行统一测试得到所述关键词的测试值。

[0115] 本发明测试设备的另一个实施例框图如图 5 所示,该测试设备包括:构造单元 510、搜索结果获取单元 520、评估单元 530 和测试单元 540。

[0116] 其中,构造单元 510 用于构造关键词集合作为测试用例的输入数据;搜索结果获取单元 520 用于根据输入的关键词获得搜索结果;评估单元 530 用于通过若干评估规则对所述搜索结果逐一进行评估,获取各个搜索结果的评估值;测试单元 540 用于通过至少一种测试算法对所述各个搜索结果的评估值进行统一测试,得到所述关键词的测试值。

[0117] 具体的,评估单元 530 包括:第一设置单元 531,用于为所述若干评估规则分别设置权重值;提取单元 532,用于顺序提取一条搜索结果;第一获取单元 533,用于分别采用

所述评估规则对所述搜索结果进行评估,得到对应每种评估规则的结果值;第一计算单元 534,用于根据所述权重值和所述结果值计算所述搜索结果的评估值。

[0118] 具体的,测试单元 540 包括:第二设置单元 541,用于有至少两种测试算法时,为所述至少两种测试算法分别设置权重值;第二获取单元 542,用于分别采用所述测试算法对所述评估值进行测试,得到对应每种测试算法的结果值;第二计算单元 543,用于根据所述权重值和所述结果值计算所述关键词的测试值。

[0119] 由以上本发明实施例可见,应用本发明对通过搜索引擎获取的测试用例进行测试时,由于采用多种测试规则的组合对搜索结果进行自动的评测打分,使得测试结果更加正确,提高了测试结果的合理性和准确性,通过该测试方式可以为后续搜索结果的排序优化提供了参考;并且通过将上述测试规则通过软件实现,可以对测试用例进行自动化测试,降低了测试过程的复杂性,提高了测试效率。

[0120] 通过以上的实施方式的描述,本领域的技术人员可以清楚地了解到本发明可借助软件加必需的通用硬件平台的方式来实现。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来,该计算机软件产品可以存储在存储介质中,如 ROM/RAM、磁碟、光盘等,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备等)执行本发明各个实施例或者实施例的某些部分所述的方法。

[0121] 以上所述的本发明实施方式,并不构成对本发明保护范围的限定。任何在本发明的精神和原则之内所作的修改、等同替换和改进等,均应包含在本发明的保护范围之内。

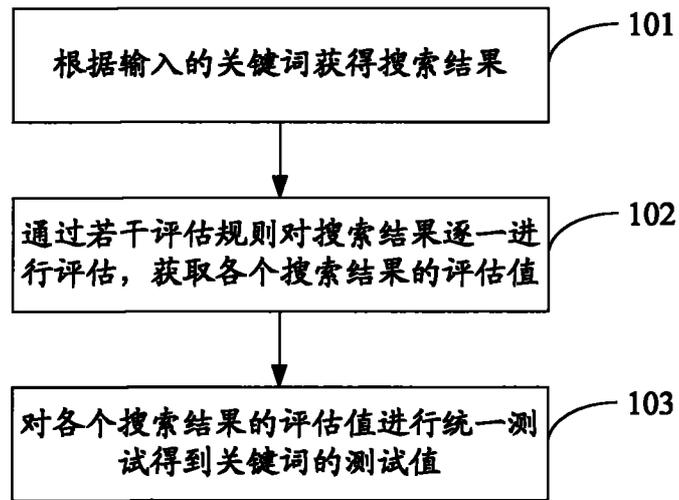


图 1

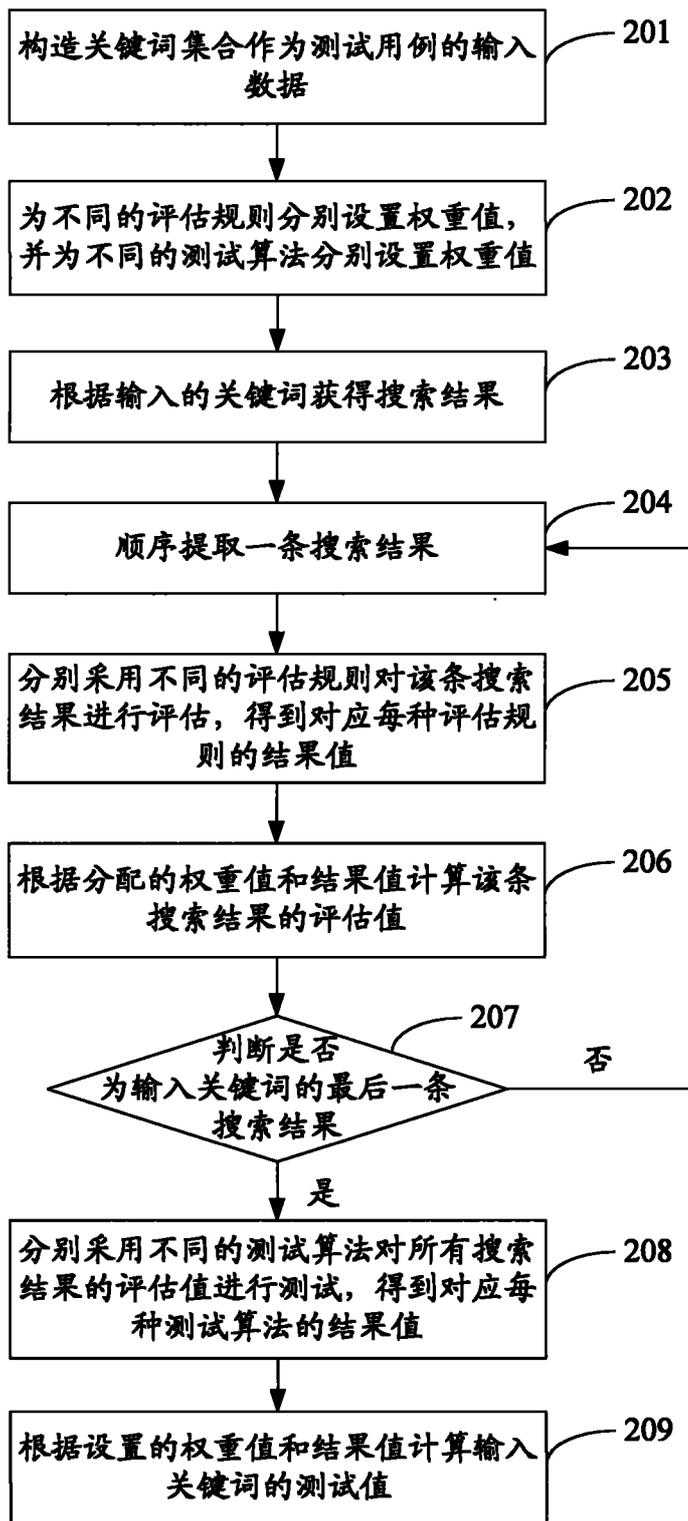


图 2

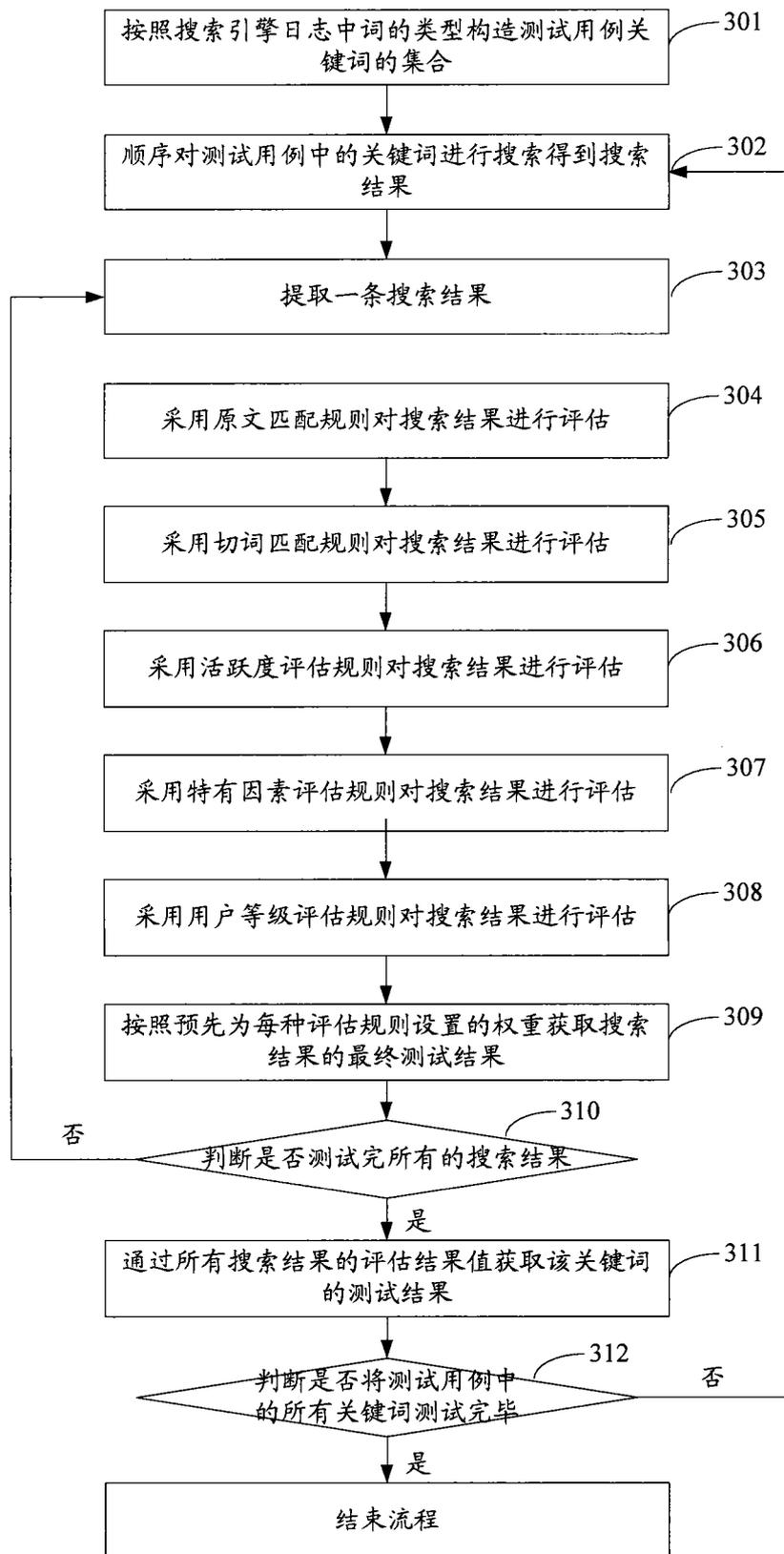


图 3

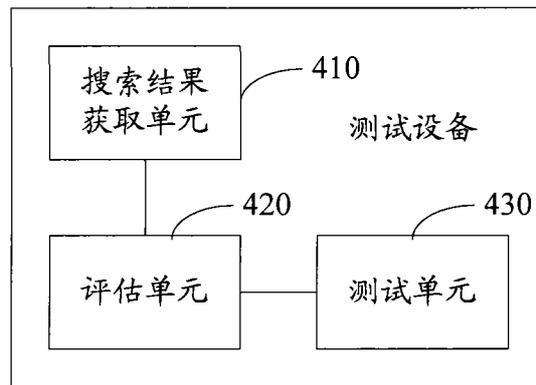


图 4

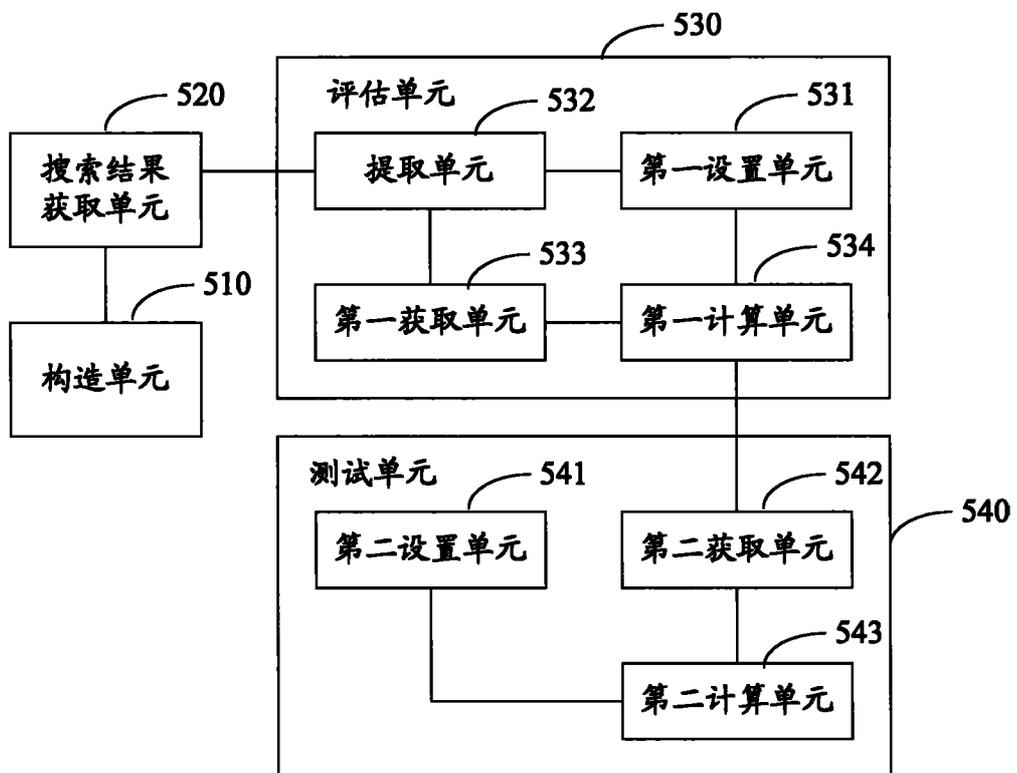


图 5