

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2015 年 11 月 26 日 (26.11.2015) WIPO | PCT



(10) 国际公布号
WO 2015/176624 A1

- (51) 国际专利分类号:
G06F 17/30 (2006.01)
- (21) 国际申请号: PCT/CN2015/079044
- (22) 国际申请日: 2015 年 5 月 15 日 (15.05.2015)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201410211458.1 2014 年 5 月 19 日 (19.05.2014) CN
201410234627.3 2014 年 5 月 29 日 (29.05.2014) CN
- (71) 申请人: 北京奇虎科技有限公司 (BEIJING QIHOO TECHNOLOGY COMPANY LIMITED) [CN/CN]; 中国北京市西城区新街口外大街 28 号 D 座 112 室 (德胜园区), Beijing 100088 (CN)。奇智软件 (北京) 有限公司 (QIZHI SOFTWARE (BEIJING) COMPANY LIMITED) [CN/CN]; 中国北京市朝阳区酒仙桥路 6 号院 2 号楼 B 座 2 层、3 层 301-306 室, Beijing 100015 (CN)。
- (72) 发明人: 苏文杰 (SU, Wenjie); 中国北京市朝阳区酒仙桥路 6 号院 2 号楼, Beijing 100015 (CN)。常富洋 (CHANG, Fuyang); 中国北京市朝阳区酒仙桥路 6 号院 2 号楼, Beijing 100015 (CN)。秦吉胜 (QIN, Jisheng); 中国北京市朝阳区酒仙桥路 6 号院 2 号楼, Beijing 100015 (CN)。周鸿祎 (ZHOU, Hongyi); 中国北京市朝阳区酒仙桥路 6 号院 2 号楼, Beijing 100015 (CN)。
- (74) 代理人: 北京智汇东方知识产权代理事务所 (普通合伙) (WISEAST INTELLECTUAL PROPERTY LAW FIRM); 中国北京市海淀区花园路 13 号 5 幢 320 房间, Beijing 100088 (CN)。
- (81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA,

[见续页]

(54) Title: SUDDEN TIMELINESS SEARCH TERM IDENTIFICATION METHOD AND SYSTEM

(54) 发明名称: 突发时效性查询词识别方法和系统

计算查询词在第一段时间内和第二段时间内的查询量变化趋势, 第二段时间为第一段时间的一部分

110

比较查询词在第一段时间和第二段时间内的查询量变化趋势之间的差距, 根据差距的大小判断查询词是否为突发时效性查询词

120

获取查询词对应的多个搜索结果条日, 多个搜索结果条日包含新闻信息条日

130

为多个搜索结果条日分别设置权值

140

在查询词为突发时效性查询词时, 增大新闻信息条日的权值

150

根据多个搜索结果条目的权值大小在查询词对应的搜索结果页上, 确定多个搜索结果条目的排名位置

160

110 Calculating the trend of change in the search quantities of a search term in a first time period and a second time period, the second time period being a part of the first time period

120 Comparing the difference between the trend of change in the search quantity of the search term in the first time period and the trend of change in the search quantity of the search term in the second time period, and according to the difference, determining whether the search term is a sudden timeliness search term

130 Acquiring a plurality of search result entries corresponding to the search term, the plurality of search result entries comprising a news information entry

140 Respectively setting a weight for the plurality of search result entries

150 When the search term is a sudden timeliness search term, increasing the weight of the news information entry

160 According to the weights of the plurality of search result entries, determining the ranking of the plurality of search result entries on a search result page corresponding to the search term

(57) Abstract: A sudden timeliness search term identification method and system, and a method and system for inserting a news information entry in a search result page, the sudden timeliness search term identification method comprising: respectively calculating the trend of change in the search quantities of a search term in a first time period and a second time period, the second time period being a part of the first time period (110); comparing the difference between the trend of change in the search quantity of the search term in the first time period and the trend of change in the search quantity of the search term in the second time period, and according to the difference, determining whether the search term is a sudden timeliness search term (120). The present invention compares the long-term trend of change with the short-term trend of change in the search quantities of a search term in a time period, and even if the search quantity of the search term is maintained at a smaller magnitude, the burst of the search quantity is obviously reflected in the difference of the trend of change in the search quantity, thus facilitating the identification of a sudden timeliness search term.

(57) 摘要:

[见续页]

图 3 / FIG.3



WO 2015/176624 A1



RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

- (84) **指定国** (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR,

本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

一种突发时效性查询词识别方法和系统, 以及一种在搜索结果页上插入新闻信息条目的方法和系统。其中, 突发时效性查询词识别方法包括: 分别计算查询词在第一段时间内和第二段时间内的查询量变化趋势, 第二段时间为第一段时间的一部分(110); 比较查询词在第一段时间和第二段时间内的查询量变化趋势之间的差距, 并根据差距的大小判断查询词是否为突发时效性查询词(120)。其比较的是查询词在一段时间内查询量的长期变化趋势与短期变化趋势, 即使查询词的查询量维持在较小的数量级, 查询量的突增情况也可以明显地体现在查询量变化趋势的差距中, 有利于识别出突发时效性查询词。

突发时效性查询词识别方法和系统

技术领域

本发明涉及计算机技术领域，特别是一种突发时效性查询词识别方法和系统，
5 以及一种在搜索结果页上插入新闻信息条目的方法和系统。

背景技术

突发时效性查询词是短时间内用户对其的查询量较之前明显变大的查询词，所以突发时效性查询词往往反映的是一个新的事件的爆发，引发用户对事件的关注，
10 其具有很强的时效性。因此，及时识别出突发时效性查询词，并根据突发时效性查询词来调整搜索引擎中新闻的排名，可以有效提升用户对新闻的搜索体验，改善搜索引擎的质量。

现有的识别突发时效性查询词的方法，是根据查询词在当前时间段的查询量和在前一段时间的查询量比较，差值越大越说明该查询词可能是当前时间段的突发时效性查询词。
15

上述技术方案的缺点在于，如两段时间内查询词的查询量维持较小的量级，则可能造成查询词的查询量出现了突增但差值却较小的情况，而差值过小则造成无法识别出该突发时效性查询词，例如：

1、一些查询量大的查询词的查询量改变会掩盖突发时效性查询词，比如，查询
20 词 A 的查询量由 10000 变为 11000，那么差值就为 1000，而突发时效性查询词的查询量由 100 变为 200，查询量虽然翻倍但差值仅为 100，与查询词 A 差值相比较小，无法体现出查询量的突增变化，造成难以识别出该突发时效性查询词。

2、一般地，统计查询量的时间段设置的比较小（如按小时设置），由于查询词的查询量受时间影响较大，在凌晨时段的整体查询量较小，在凌晨比较查询量得到的差值也较小，所以难以识别出突发时效性查询词；如为了克服该问题，将时间段设置的比较大（如按天设置），进行不同时间段的查询量比较的时间间隔就比较久，那么突发时效性查询词会更新比较慢。
25

另外，搜索引擎在接收到用户输入的查询词之后，会返回给用户与查询词相关的多个搜索结果条目，该多个搜索结果条目会按顺序显示在搜索结果页上。

30 现有的对于搜索结果条目排序的方式，普遍是将较旧网页（例如被点击次数较多的网页）对应的搜索结果条目放在较靠前的位置。这种排序方式的缺陷在于：对于新闻信息网页来说，由于普遍为较新的网页，所以其对应的新闻信息条目无法展现在搜索结果页上较前的位置；而用户输入查询词的目的经常是为了搜索查询词相关的突发或热门事件，则现有的对搜索结果条目进行排序的方案会导致用户无法及
35 时发现记录有查询词相关的突发或热门事件的新闻网页。

由此可见，现有的对搜索结果条目排序的方式由于不能识别出突发或热门事件对应的网页，所以难以恰当地设置突发或热门事件对应的新闻信息网页的条目在搜索结果页上的位置，导致无法将突发或热门事件对应的网页及时推送给用户。

所以，本发明需要解决的技术问题在于，提供一种新的识别突发时效性查询词的技术方案，保证即使在查询词的查询量始终维持在较小量级的情况下，也能有效地识别出突发时效性查询词。以及，进一步如何根据突发时效性查询词来调整搜索引擎中新闻网页的排名，从而将突发或热门事件对应的网页及时推送给用户。

5

发明内容

鉴于上述问题，提出了本发明以便提供一种克服上述问题或者至少部分地解决上述问题的突发时效性查询词识别方法和系统，以及在搜索结果页上插入新闻信息条目的方法和系统。

10 根据本发明的一方面，提供了一种突发时效性查询词识别方法，其包括：分别计算查询词在第一段时间内和第二段时间内的查询量变化趋势，所述第二段时间为所述第一段时间的一部分；比较所述查询词在所述第一段时间和所述第二段时间内的查询量变化趋势之间的差距，并根据所述差距的大小判断所述查询词是否为突发时效性查询词。

15 根据本发明的另一方面，提供了一种突发时效性查询词识别系统，其包括：变化趋势计算模块，用于分别计算查询词在第一段时间内和第二段时间内的查询量变化趋势，所述第二段时间为所述第一段时间的一部分；识别模块，用于比较所述查询词在所述第一段时间和所述第二段时间内的查询量变化趋势之间的差距，并根据所述差距的大小判断所述查询词是否为突发时效性查询词。

20 根据以上技术方案，本发明的突发时效性查询词识别方法和系统至少有以下优点：由于第二段时间为第一段时间的一部分，所以本发明的技术方案中比较的是查询词在一段时间内查询量的长期变化趋势与短期变化趋势，如果第一段时间内查询词的查询量变化趋势表现为缓慢增加，而在第二段时间内查询词的查询量变化趋势表现为急速增加，则变化趋势的差距就说明突发地出现大量用户关注该查询词的相关新闻，并在搜索引擎中进行查询，则该查询词必然是突发时效性查询词；查询词的查询量变化趋势与其查询量无关，即使查询词的查询量维持在较小的数量级，查询量的突增情况也可以明显地体现在查询量变化趋势的差距中，有利于识别出突发时效性查询词，而不受到较大查询量的查询词干扰，不受到整体查询量较小的时间段影响。

30

根据本发明的又一方面，提供了一种在搜索结果页上插入新闻信息条目的方法，其包括：接收用户的查询词，并根据所述查询词获取相关的实体词；获取所述实体词的第一事件属性特征；获取所述查询词对应的新闻信息条目；计算所述新闻信息条目的第二事件属性特征；将所述第一事件属性特征与所述第二事件属性特征进行
35 比较，根据比较结果计算所述查询词的事件时效值；根据所述查询词的事件时效值的大小，确定所述新闻信息条目插入到所述查询词对应的搜索结果页上的排名位置。

根据本发明的又一方面，提供了一种在搜索结果页上插入新闻信息条目的系统，其包括：实体词获取模块，用于接收用户的查询词，并根据所述查询词获取相关的

实体词；第一事件属性特征获取模块，用于获取所述实体词的第一事件属性特征；新闻信息条目获取模块，用于获取所述查询词对应的新闻信息条目；第二事件属性特征计算模块，用于计算所述新闻信息条目的第二事件属性特征；事件时效值计算模块，用于将所述第一事件属性特征与所述第二事件属性特征进行比较，根据比较结果计算所述查询词的事件时效值；排名位置确定模块，用于根据所述查询词的事件时效值的大小，确定所述新闻信息条目插入到所述查询词对应的搜索结果页上的排名位置。

根据以上技术方案，本发明的在搜索结果页上插入新闻信息条目的方法和系统至少具有以下优点：第一事件属性特征是与查询词中实体词相关的已收集新闻类消息中事件的特征，第二事件属性特征是与查询词相关的特定新闻类消息中事件的特征，所以将第一事件属性特征与第二事件属性特征进行比较得到的时效值，反映了特定新闻类消息相对于已收集新闻类消息的时效特点，即突发或热门程度，根据事件时效值高低来控制新闻信息条目排名的话，可以保证突发或热门事件对应的新闻类消息可以及时被用户发现，从而提高用户的使用体验。

根据本发明的又一方面，提供了一种计算机程序，其包括计算机可读代码，当所述计算机可读代码在计算设备上运行时，导致所述计算设备执行根据上文所述的突发时效性查询词识别方法和/或在搜索结果页上插入新闻信息条目的方法。

根据本发明的再一方面，提供了一种计算机可读介质，其中存储了上述的计算机程序。

上述说明仅是本发明技术方案的概述，为了能够更清楚了解本发明的技术手段，而可依照说明书的内容予以实施，并且为了让本发明的上述和其它目的、特征和优点能够更明显易懂，以下特举本发明的具体实施方式。

附图说明

通过阅读下文优选实施方式的详细描述，各种其他的优点和益处对于本领域普通技术人员将变得清楚明了。附图仅用于示出优选实施方式的目的，而并不认为是对本发明的限制。而且在整个附图中，用相同的参考符号表示相同的部件。在附图中：

图 1 示出了根据本发明的一个实施例的突发时效性查询词识别方法的流程图；

图 2 示出了根据本发明的另一个实施例的突发时效性查询词识别方法的流程图；

图 3 示出了根据本发明的又一个实施例的突发时效性查询词识别方法的流程图；

图 4 示出了根据本发明的一个实施例的突发时效性查询词识别系统的框图；

图 5 示出了根据本发明的另一个实施例的突发时效性查询词识别系统的框图；

图 6 示出了根据本发明的又一个实施例的突发时效性查询词识别方法的框图；

图 7 示出了根据本发明的一个实施例的在搜索结果页上插入新闻信息条目的方法的流程图；

图 8 示出了根据本发明的一个实施例的在搜索结果页上插入新闻信息条目的系统的框图；

图 9 示意性地示出了用于执行根据本发明的突发时效性查询词识别方法和/或在搜索结果页上插入新闻信息条目的方法的计算设备的框图；以及

5 图 10 示意性地示出了用于保持或者携带实现根据本发明的突发时效性查询词识别方法和/或在搜索结果页上插入新闻信息条目的方法的程序代码的存储单元。

具体实施方式

下面结合附图和具体的实施方式对本发明作进一步的描述。

10 如图 1 所示，本发明的一个实施例中提供了一种突发时效性查询词识别方法，其包括：

步骤 110，分别计算查询词在第一段时间内和第二段时间内的查询量变化趋势，第二段时间为第一段时间的一部分。在本实施例的技术方案中，对查询量变化趋势的表现形式不做限制，例如，可以取多个时间点，记录截止到每个时间点的总查询
15 量，则多个时间点对应的总查询量可提现查询量变化趋势。

步骤 120，比较查询词在第一段时间和第二段时间内的查询量变化趋势之间的差距，并根据差距的大小判断查询词是否为突发时效性查询词。在本实施例的技术方案中，对查询量变化趋势的差距大小的比较形式不做限制，例如，可以分别计算第一段时间内查询词的查询量在单位时间内的变化量，以及第二段时间内查询词的
20 查询量在单位时间内的变化量，并进行比较。

在本实施例的技术方案中，由于第二段时间为第一段时间的一部分，所以本发明的技术方案中比较的是查询词在一段时间内查询量的长期变化趋势与短期变化趋势，如果第一段时间内查询词的查询量变化趋势表现为缓慢增加，而在第二段时间内查询词的查询量变化趋势表现为急速增加，则变化趋势的差距就说明突发地出现
25 大量用户关注该查询词的相关新闻，并在搜索引擎中进行查询，则该查询词必然是突发时效性查询词。查询词的查询量变化趋势与其查询量无关，即使查询词的查询量维持在较小的数量级，查询量的突增情况也可以明显地体现在查询量变化趋势的差距中，有利于识别出突发时效性查询词，而不受到较大查询量的查询词干扰，不受到整体查询量较小的时间段影响。

30 如图 2 所示，本发明的另一实施例中还提供了一种突发时效性查询词识别方法，相比于上述实施例，本实施例的突发时效性查询词识别方法，步骤 110 包括：

步骤 111，按预设长度的时间段，连续多次统计查询词的查询量。本实施例中
对时间段的大小不做限制，为了及时识别出突发时效性查询词，可以将时间段设置的较小，如按小时。

35 步骤 112，取最近的连续 m 个时间段为第一段时间，并根据查询词在 m 个时间段内的查询量，确定查询词在第一段时间内的查询量变化趋势， m 为整数。

步骤 113，取最近的连续 n 个时间段为第二段时间，并根据查询词在 n 个时间段内的查询量，确定查询词在第二段时间内的查询量变化趋势， n 为小于 m 的整数。

本实施例中对 n 和 m 的具体大小不做限制，为了突出第一段时间为长期，第二段时间为短期， m 的值可以远大于 n ，例如时间段为 1 小时的话，则 m 可以取 48， n 可以取 5，即第一段时间为 48 小时，第二段时间为 5 小时。

在本实施例的技术方案中，不需针对两段时间内分别统计查询量再计算查询量变化趋势，而只需按时间段连续地统计一次查询量即可，只要分别对应两端时间的多个时间段分别计算查询量变化趋势即可。

本发明的另一实施例中还提供了一种突发时效性查询词识别方法，相比于上述实施例，本实施例的突发时效性查询词识别方法，步骤 120 包括：

取反映查询词在第一段时间内的查询量变化趋势的第一轨迹线，以及反映查询词在第二段时间内的查询量变化趋势的第二轨迹线，并将第一轨迹线以及第二轨迹线的夹角作为差距。在本实施例中，对于轨迹线的类型不做限制，以横坐标为时间段、纵坐标为查询量，多个时间段的查询量反映为多个点，则例如轨迹线可以是连接多个时间上的查询量的折线，也可以是通过最小二乘法拟合的平滑曲线。

在本实施例的技术方案中，轨迹线可以直观准确地体现查询量变化趋势，而轨迹线夹角大小可以直观准确地体现查询量变化趋势的差距，则本实施例的技术方案查询量变化趋势的差距进行量化，量化的结果有利于准确识别查询词是否为突发时效性查询词。

如图 3 所示，本发明的又一实施例中还提供了一种突发时效性查询词识别方法，相比于上述实施例，本实施例的突发时效性查询词识别方法，还包括：

步骤 130，获取查询词对应的多个搜索结果条目，多个搜索结果条目中包含新闻信息条目。

步骤 140，为多个搜索结果条目分别设置权值。

步骤 150，在查询词为突发时效性查询词时，增大新闻信息条目的权值。

步骤 160，根据多个搜索结果条目的权值大小，在查询词对应的搜索结果页上，确定多个搜索结果条目的排名位置。一般地，权值越大的搜索结果条目，排名位置越高。

在本实施例的技术方案中，对于突发时效性查询词来说，提高了其相关新闻条目的排名位置，有利于用户及时发现其关注的突发性事件，提升了用户体验。

如图 4 所示，本发明的一个实施例提供了一种突发时效性查询词识别系统，其包括：

变化趋势计算模块 410，用于分别计算查询词在第一段时间内和第二段时间内的查询量变化趋势，第二段时间为第一段时间的一部分。在本实施例的技术方案中，对查询量变化趋势的表现形式不做限制，例如，可以取多个时间点，记录截止到每个时间点的总查询量，则多个时间点对应的总查询量可提现查询量变化趋势。

识别模块 420，用于比较查询词在第一段时间和第二段时间内的查询量变化趋势之间的差距，并根据差距的大小判断查询词是否为突发时效性查询词。在本实施例的技术方案中，对查询量变化趋势的差距大小的比较形式不做限制，例如，可以分别计算第一段时间内查询词的查询量在单位时间的内变化量，以及第二段时间内查询词的查询量在单位时间的内变化量，并进行比较。

在本实施例的技术方案中，由于第二段时间为第一段时间的一部分，所以本发明的技术方案中比较的是查询词在一段时间内查询量的长期变化趋势与短期变化趋势，如果第一段时间内查询词的查询量变化趋势表现为缓慢增加，而在第二段时间内查询词的查询量变化趋势表现为急速增加，则变化趋势的差距就说明突发地出现大量用户关注该查询词的相关新闻，并在搜索引擎中进行查询，则该查询词必然是突发时效性查询词。查询词的查询量变化趋势与其查询量无关，即使查询词的查询量维持在较小的数量级，查询量的突增情况也可以明显地体现在查询量变化趋势的差距中，有利于识别出突发时效性查询词，而不受到较大查询量的查询词干扰，不受到整体查询量较小的时间段影响。

如图 5 所示，本发明的另一实施例中还提供了一种突发时效性查询词识别系统，相比于上述实施例，本实施例的突发时效性查询词识别系统，还包括：

查询量统计模块 430，用于按预设长度的时间段，连续多次统计查询词的查询量。本实施例中对时间段的大小不做限制，为了及时识别出突发时效性查询词，可以将时间段设置的较小，如按小时。

变化趋势计算模块 410 取最近的连续 m 个时间段为第一段时间，并根据查询词在 m 个时间段内的查询量，确定查询词在第一段时间内的查询量变化趋势， m 为整数。

变化趋势计算模块 410 取最近的连续 n 个时间段为第二段时间，并根据查询词在 n 个时间段内的查询量，确定查询词在第二段时间内的查询量变化趋势， n 为小于 m 的整数。

本实施例中对 n 和 m 的具体大小不做限制，为了突出第一段时间为长期，第二段时间为短期， m 的值可以远大于 n ，例如时间段为 1 小时的话，则 m 可以取 48， n 可以取 5，即第一段时间为 48 小时，第二段时间为 5 小时。

在本实施例的技术方案中，不需针对两段时间内分别统计查询量再计算查询量变化趋势，而只需按时间段连续地统计一次查询量即可，只要分别对应两端时间的多个时间段分别计算查询量变化趋势即可。

本发明的另一实施例中还提供了一种突发时效性查询词识别系统，相比于上述实施例，本实施例的突发时效性查询词识别系统，识别模块 420 取反映查询词在第一段时间内的查询量变化趋势的第一轨迹线，以及反映查询词在第二段时间内的查询量变化趋势的第二轨迹线，并将第一轨迹线以及第二轨迹线的夹角作为差距。在本实施例中，对于轨迹线的类型不做限制，以横坐标为时间段、纵坐标为查询量，多个时间段的查询量反映为多个点，则例如轨迹线可以是连接多个时间上的查询量的折线，也可以是通过最小二乘法拟合的平滑曲线。

在本实施例的技术方案中，轨迹线可以直观准确地体现查询量变化趋势，而轨迹线夹角大小可以直观准确地体现查询量变化趋势的差距，则本实施例的技术方案查询量变化趋势的差距进行量化，量化的结果有利于准确识别查询词是否为突发时效性查询词。

如图 6 所示，本发明的又一实施例中还提供了一种突发时效性查询词识别系统，相比于上述实施例，本实施例的突发时效性查询词识别系统，还包括：

搜索结果条目获取模块 440，用于获取查询词对应的多个搜索结果条目，多个搜索结果条目中包含新闻信息条目。

权值设置模块 450，用于为多个搜索结果条目分别设置权值。

5 权值调整模块 460，用于在查询词为突发时效性查询词时，增大新闻信息条目的权值。

排名位置确定模块 470，用于根据多个搜索结果条目的权值大小，在查询词对应的搜索结果页上，确定多个搜索结果条目的排名位置。一般地，权值越大的搜索结果条目，排名位置越高。

10 在本实施例的技术方案中，对于突发时效性查询词来说，提高了其相关新闻条目的排名位置，有利于用户及时发现其关注的突发性事件，提升了用户体验。

根据本发明实施例提供，本发明实施例还提供一种在搜索结果页上插入新闻信息条目的方法和系统。

15 如图 7 所示，本发明的一个实施例提供了一种在搜索结果页上插入新闻信息条目的方法，其包括：

步骤 710，接收用户的查询词，并根据查询词获取相关的实体词。实体词是通过预先分析反映事件的新闻类消息（页面或实时性消息）的标题/主题，进行分词技术后提取出的能代表事件的整体或部分特征的名词或动词。可选地，在判断查询词中不包含有实体词或扩展不到实体词时，则可以确定用户输入查询词并非是为了查
20 询某事件，可以不执行后续的操作。

步骤 720，获取实体词的第一事件属性特征。由于事件由新闻类消息发布，而实体词反映了事件的特征，所以实体词的第一事件属性特征可以根据已收集的含有该实体词的新闻类消息得到，则第一事件属性特征包括但不限于：新闻类消息中实体词出现的次数/频率、具有实体词的新闻类信息的发布日期、一段时间内有关实体
25 词的新闻类消息的频率，或一段时间该实体词相关的新闻类消息的转载次数等。

步骤 730，获取查询词对应的新闻信息条目。本实施例中对新闻信息条目的来源不做限定，例如可以是新闻站点推送的最新新闻消息对应的条目，也可以是最新抓取到的新闻消息类网页，或者社交网站上的新闻消息等等。

30 步骤 740，计算新闻信息条目的第二事件属性特征。与第一事件属性特征相似地，第二事件属性特征包括但不限于：新闻信息条目对应的新闻类消息中实体词出现的次数/频率、新闻信息条目对应的新闻类消息的发布日期、新闻信息条目对应的新闻类消息的出现频率，或一段时间该则新闻信息条目对应的新闻类消息的转载次数等。

35 步骤 750，将第一事件属性特征与第二事件属性特征进行比较，根据比较结果计算查询词的事件时效值。通过采用不同方式进行第一事件属性特征和第二事件属性特征之间的比较所得到的事件时效值，体现了与查询词相关的实体词的新闻热度，与查询词本身相关到的新闻类消息的新闻热度之间的关联性，也体现新闻信息条目所对应的新闻类信息与已收集的新闻类信息之间的差异、相似度或其他能体现时效性的关系，这些关系能够反映出：相对于离线时已收集的新闻类信息，查询词所关

联的新闻信息条目所对应的新闻类信息中事件的突发或热门程度，越热门则后续给予插入的新闻条目的搜索结果排名应该越靠前。

步骤 760，根据查询词的事件时效值的大小，确定新闻信息条目插入到查询词对应的搜索结果页上的排名位置。则根据本实施例的技术方案，实际上是根据新闻信息条目对应新闻类消息中的事件的突发或热门程度高低，调整新闻信息条目置于搜索结果页中的排名高低，突发或热门程度高则排名靠前，有利于将突发或热门程度高的事件的相关新闻类消息及时推送给用户。一般地，搜索引擎会为每个搜索结果条目提供一个排名值，则本实施例中可以根据事件时效值大小为新闻信息条目或将条目集合成一条搜索结果条目并赋予排名值，以进行在搜索结果页上的排名。

本发明的另一实施例提出一种在搜索结果页上插入新闻信息条目的方法，与上述实施例相比，本实施例的在搜索结果页上插入新闻信息条目的方法，步骤 710 包括：从查询词中提取实体词。本实施例中对于提取实体词的方式不做限定，例如，可以直接使用现有搜索引擎的分词技术进行提取。

本发明的另一实施例提出一种在搜索结果页上插入新闻信息条目的方法，与上述实施例相比，本实施例的在搜索结果页上插入新闻信息条目的方法，步骤 710 包括：根据查询词进行扩展得到相关的实体词。本实施例中对于扩展实体词的方式不做限定，例如，可以选择与查询词同步率（用户同时输入进行搜索的频率）较高的词语对查询词进行扩展，例如，近期内与查询词“水灾”同步率较高的词为“A市”，则将查询词“水灾”扩展为“水灾 A市”。

本发明的另一实施例提出一种在搜索结果页上插入新闻信息条目的方法，与上述实施例相比，本实施例的在搜索结果页上插入新闻信息条目的方法，在步骤 720 之前，还包括：

抓取新闻信息页面，并根据包含实体词的新闻信息页面计算第一事件属性特征。本实施例的技术方案，预先抓取新闻信息页面以用于分析第一事件属性特征。可选地，为了降低第一事件属性特征的计算量，可以仅选择标题中包含实体词的新闻信息页面进行计算。且第一事件属性特征、实体词之间可按照对应关系建立索引，以便于按实体词查询对应的第一事件属性特征。

本发明的另一实施例提出一种在搜索结果页上插入新闻信息条目的方法，与上述实施例相比，本实施例的在搜索结果页上插入新闻信息条目的方法，步骤 740 包括：

根据新闻信息条目对应的页面，计算第二事件属性特征。新闻信息条目对应页面中的内容，即为相应的新闻类消息的事件具体内容，可以用于计算第二事件属性特征。

本发明的另一实施例提出一种在搜索结果页上插入新闻信息条目的方法，与上述实施例相比，本实施例的在搜索结果页上插入新闻信息条目的方法，步骤 730 包括：

从已抓取的新闻信息页面中查询出查询词对应的新闻信息页面，并根据查询词对应的新闻信息页面设置新闻信息条目。最新抓取的新闻信息页面往往未经搜索引擎进行整理，则用户通过搜索引擎往往搜索不到该页面对应的条目，所以通过本实

施例的技术方案，利于将突发或热门事件的最新页面及时推送给用户。

本发明的另一实施例提出一种在搜索结果页上插入新闻信息条目的方法，与上述实施例相比，本实施例的在搜索结果页上插入新闻信息条目的方法，步骤 730 包括：

- 5 通过搜索引擎搜索查询词对应的多个搜索结果条目，并从多个搜索结果条目中提取出新闻信息条目。根据本实施例的技术方案，如果新闻信息条目对应新闻类消息中的事件突发或热门程度高，则可以提到较高排名位置进行显示，利于将突发或热门事件的最新新闻类消息及时推送给用户。

10 如图 8 所示，本发明的另一实施例提供了一种在搜索结果页上插入新闻信息条目的系统，其包括：

 实体词获取模块 810，用于接收用户的查询词，并根据查询词获取相关的实体词。实体词是通过预先分析反映事件的新闻类消息（页面或实时性消息）的标题/主题，进行分词技术后提取出的能代表事件的整体或部分特征的名词或动词。可选地，在判断查询词中不包含有实体词或扩展不到实体词时，则可以确定用户输入查询词并非是为了查询某事件，可以不执行后续的操作。

15 第一事件属性特征获取模块 820，用于获取实体词的第一事件属性特征。由于事件由新闻类消息发布，而实体词反映了事件的特征，所以实体词的第一事件属性特征可以根据已收集的含有该实体词的新闻类消息得到，则第一事件属性特征包括但不限于：新闻类消息中实体词出现的次数/频率、具有实体词的新闻类信息的发布日期、一段时间内有关实体词的新闻类消息的频率，或一段时间该实体词相关的新闻类消息的转载次数等。

 新闻信息条目获取模块 830，用于获取查询词对应的新闻信息条目。本实施例中新闻信息条目的来源不做限定，例如可以是新闻站点推送的最新新闻消息对应的条目，也可以是最新抓取到的新闻消息类网页，或者社交网站上的新闻消息等等。

25 第二事件属性特征计算模块 840，用于计算新闻信息条目的第二事件属性特征。与第一事件属性特征相似地，第二事件属性特征包括但不限于：新闻信息条目对应的新闻类消息中实体词出现的次数/频率、新闻信息条目对应的新闻类消息的发布日期、新闻信息条目对应的新闻类消息的出现频率，或一段时间该则新闻信息条目对应的新闻类消息的转载次数等。

30 事件时效值计算模块 850，用于将第一事件属性特征与第二事件属性特征进行比较，根据比较结果计算查询词的事件时效值。通过采用不同方式进行第一事件属性特征和第二事件属性特征之间的比较所得到的事件时效值，体现了与查询词相关的实体词的新闻热度，与查询词本身相关到的新闻类消息的新闻热度之间的关联性，也体现新闻信息条目所对应的新闻类信息与已收集的新闻类信息之间的差异、相似
35 度或其他能体现时效性的关系，这些关系能够反映出：相对于离线时已收集的新闻类信息，查询词所关联的新闻信息条目所对应的新闻类信息中事件的突发或热门程度，越热门则后续给予插入的新闻条目的搜索结果排名应该越靠前。

 排名位置确定模块 860，用于根据查询词的事件时效值的大小，确定新闻信息条目插入到查询词对应的搜索结果页上的排名位置。则根据本实施例的技术方案，

实际上是根据新闻信息条目对应新闻类消息中的事件的突发或热门程度高低，调整新闻信息条目置于搜索结果页中的排名高低，突发或热门程度高则排名靠前，有利于将突发或热门程度高的事件的相关新闻类消息及时推送给用户。一般地，搜索引擎会为每个搜索结果条目提供一个排名值，则本实施例中可以根据事件时效值大小

5 为新闻信息条目或将条目集成一条搜索结果条目并赋予排名值，以进行在搜索结果页上的排名。

本发明的另一实施例提出一种在搜索结果页上插入新闻信息条目的系统，与上述实施例相比，本实施例的在搜索结果页上插入新闻信息条目的系统，实体词获取模块 810 从查询词中提取实体词。本实施例中对于提取实体词的方式不做限定，例如，可以直接使用现有搜索引擎的分词技术进行提取。

10

本发明的另一实施例提出一种在搜索结果页上插入新闻信息条目的系统，与上述实施例相比，本实施例的在搜索结果页上插入新闻信息条目的系统，实体词获取模块 810 根据查询词进行扩展得到相关的实体词。本实施例中对于扩展实体词的方式不做限定，例如，可以选择与查询词同步率（用户同时输入进行搜索的频率）较高的词语对查询词进行扩展，例如，近期内与查询词“水灾”同步率较高的词为“A市”，则将查询词“水灾”扩展为“水灾 A 市”。

15

本发明的另一实施例提出一种在搜索结果页上插入新闻信息条目的系统，与上述实施例相比，本实施例的在搜索结果页上插入新闻信息条目的系统，第一事件属性获取模块 820 抓取新闻信息页面，并根据包含实体词的新闻信息页面，计算第一事件属性特征。本实施例的技术方案，预先抓取新闻信息页面以用于分析第一事件属性特征。可选地，为了降低第一事件属性特征的计算量，可以仅选择标题中包含实体词的新闻信息页面进行计算。且第一事件属性特征、实体词之间可按照对应关系建立索引，以便于按实体词查询对应的第一事件属性特征。

20

本发明的另一实施例提出一种在搜索结果页上插入新闻信息条目的系统，与上述实施例相比，本实施例的在搜索结果页上插入新闻信息条目的系统，第二事件属性特征计算模块 840 根据新闻信息条目对应的页面，计算第二事件属性特征。新闻信息条目对应页面中的内容，即为相应的新闻类消息的事件具体内容，可以用于计算第二事件属性特征。

25

本发明的另一实施例提出一种在搜索结果页上插入新闻信息条目的系统，与上述实施例相比，本实施例的在搜索结果页上插入新闻信息条目的系统，新闻信息条目获取模块 830 从已抓取的新闻信息页面中查询出查询词对应的新闻信息页面，并根据查询词对应的新闻信息页面设置新闻信息条目。最新抓取的新闻信息页面往往未经搜索引擎进行整理，则用户通过搜索引擎往往搜索不到该页面对应的条目，所以通过本实施例的技术方案，利于将突发或热门事件的最新页面及时推送给用户。

30

本发明的另一实施例提出一种在搜索结果页上插入新闻信息条目的系统，与上述实施例相比，本实施例的在搜索结果页上插入新闻信息条目的系统，新闻信息条目获取模块 830 通过搜索引擎搜索查询词对应的多个搜索结果条目，并从多个搜索结果条目中提取出新闻信息条目。根据本实施例的技术方案，如果新闻信息条目对应新闻类消息中的事件突发或热门程度高，则可以提到较高排名位置进行显示，利

35

于将突发或热门事件的最新新闻类消息及时推送给用户。

在此处所提供的说明书中，说明了大量具体细节。然而，能够理解，本发明的实施例可以在没有这些具体细节的情况下实践。在一些实例中，并未详细示出公知的方法、结构和技术，以便不模糊对本说明书的理解。

类似地，应当理解，为了精简本公开并帮助理解各个发明方面中的一个或多个，在上面对本发明的示例性实施例的描述中，本发明的各个特征有时被一起分组到单个实施例、图、或者对其的描述中。然而，并不应将该公开的方法解释成反映如下意图：即所要求保护的本发明要求比在每个权利要求中所明确记载的特征更多的特征。更确切地说，如下面的权利要求书所反映的那样，发明方面在于少于前面公开的单个实施例的所有特征。因此，遵循具体实施方式的权利要求书由此明确地并入该具体实施方式，其中每个权利要求本身都作为本发明的单独实施例。

本领域那些技术人员可以理解，可以对实施例中的设备中的模块进行自适应性地改变并且把它们设置在与该实施例不同的一个或多个设备中。可以把实施例中的模块或单元或组件组合成一个模块或单元或组件，以及此外可以把它分成多个子模块或子单元或子组件。除了这样的特征和/或过程或者单元中的至少一些是相互排斥之外，可以采用任何组合对本说明书（包括伴随的权利要求、摘要和附图）中公开的所有特征以及如此公开的任何方法或者设备的所有过程或单元进行组合。除非另外明确陈述，本说明书（包括伴随的权利要求、摘要和附图）中公开的每个特征可以由提供相同、等同或相似目的的替代特征来代替。

此外，本领域的技术人员能够理解，尽管在此所述的一些实施例包括其它实施例中有所包括的某些特征而不是其它特征，但是不同实施例的特征的组合意味着处于本发明的范围之内并且形成不同的实施例。例如，在下面的权利要求书中，所要求保护的实施例的任意之一都可以以任意的组合方式来使用。

本发明的各个部件实施例可以以硬件实现，或者以在一个或者多个处理器上运行的软件模块实现，或者以它们的组合实现。本领域的技术人员应当理解，可以在实践中使用微处理器或者数字信号处理器（DSP）来实现根据本发明实施例的突发时效性查询词识别系统，以及在搜索结果页上插入新闻信息条目的系统中的一些或者全部部件的一些或者全部功能。本发明还可以实现为用于执行这里所描述的方法的一部分或者全部的设备或者装置程序（例如，计算机程序和计算机程序产品）。这样的实现本发明的程序可以存储在计算机可读介质上，或者可以具有一个或者多个信号的形式。这样的信号可以从因特网网站上下载得到，或者在载体信号上提供，或者以任何其他形式提供。

例如，图 9 示出了可以实现在智能终端之间传输数据的方法的计算设备。该计算设备传统上包括处理器 910 和以存储器 920 形式的计算机程序产品或者计算机可读介质。存储器 920 可以是诸如闪存、EEPROM（电可擦除可编程只读存储器）、EPROM、硬盘或者 ROM 之类的电子存储器。存储器 920 具有用于执行上述方法中的任何方法步骤的程序代码 931 的存储空间 930。例如，用于程序代码的存储空间 930 可以包括分别用于实现上面的方法中的各种步骤的各个程序代码 931。这些程序

代码可以从一个或者多个计算机程序产品中读出或者写入到这一个或者多个计算机程序产品中。这些计算机程序产品包括诸如硬盘，紧致盘（CD）、存储卡或者软盘之类的程序代码载体。这样的计算机程序产品通常为如参考图 10 所述的便携式或者固定存储单元。该存储单元可以具有与图 9 的计算设备中的存储器 920 类似布置的存储段、存储空间等。程序代码可以例如以适当形式进行压缩。通常，存储单元包括计算机可读代码 931'，即可以由例如诸如 910 之类的处理器读取的代码，这些代码当由计算设备运行时，导致该计算设备执行上面所描述的方法中的各个步骤。

本文中所称的“一个实施例”、“实施例”或者“一个或者多个实施例”意味着，结合实施例描述的特定特征、结构或者特性包括在本发明的至少一个实施例中。

此外，请注意，这里“在一个实施例中”的词语例子不一定全指同一个实施例。

应该注意的是上述实施例对本发明进行说明而不是对本发明进行限制，并且本领域技术人员在不脱离所附权利要求的范围的情况下可设计出替换实施例。在权利要求中，不应将位于括号之间的任何参考符号构造成对权利要求的限制。单词“包含”不排除存在未列在权利要求中的元件或步骤。位于元件之前的单词“一”或“一个”不排除存在多个这样的元件。本发明可以借助于包括有若干不同元件的硬件以及借助于适当编程的计算机来实现。在列举了若干装置的单元权利要求中，这些装置中的若干个可以是通过同一个硬件项来具体体现。单词第一、第二、以及第三等的使用不表示任何顺序。可将这些单词解释为名称。

此外，还应当注意，本说明书中使用的语言主要是为了可读性和教导的目的而选择的，而不是为了解释或者限定本发明的主题而选择的。因此，在不偏离所附权利要求书的范围和精神的情况下，对于本技术领域的普通技术人员来说许多修改和变更都是显而易见的。对于本发明的范围，对本发明所做的公开是说明性的，而非限制性的，本发明的范围由所附权利要求书限定。

权 利 要 求

1、一种突发时效性查询词识别方法，其包括：

5 分别计算查询词在第一段时间内和第二段时间内的查询量变化趋势，所述第二段时间为所述第一段时间的一部分；

比较所述查询词在所述第一段时间和所述第二段时间内的查询量变化趋势之间的差距，并根据所述差距的大小判断所述查询词是否为突发时效性查询词。

2、根据权利要求1所述的突发时效性查询词识别方法，其中，所述分别计算查询词在第一段时间内和第二段时间内的查询量变化趋势的步骤包括：

10 按预设长度的时间段，连续多次统计所述查询词的查询量；

取最近的连续 m 个时间段为所述第一段时间，并根据所述查询词在所述 m 个时间段内的查询量，确定所述查询词在所述第一段时间内的查询量变化趋势， m 为整数；

15 取最近的连续 n 个时间段为所述第二段时间，并根据所述查询词在所述 n 个时间段内的查询量，确定所述查询词在所述第二段时间内的查询量变化趋势， n 为小于 m 的整数。

3、根据权利要求1所述的突发时效性查询词识别方法，其中，所述比较所述查询词在所述第一段时间和所述第二段时间内的查询量变化趋势之间的差距的步骤包括：

20 取反映所述查询词在所述第一段时间内的查询量变化趋势的第一轨迹线，以及反映所述查询词在所述第二段时间内的查询量变化趋势的第二轨迹线，并将所述第一轨迹线以及所述第二轨迹线的夹角作为所述差距。

4、根据权利要求1至3中任一项所述的突发时效性查询词识别方法，其中，还包括：

25 获取所述查询词对应的多个搜索结果条目，所述多个搜索结果条目中包含新闻信息条目；

为所述多个搜索结果条目分别设置权值；

在所述查询词为突发时效性查询词时，增大所述新闻信息条目的权值；

根据所述多个搜索结果条目的权值大小，在所述查询词对应的搜索结果页上，

30 确定所述多个搜索结果条目的排名位置。

5、一种在搜索结果页上插入新闻信息条目的方法，其包括：

接收用户的查询词，并根据所述查询词获取相关的实体词；

获取所述实体词的第一事件属性特征；

获取所述查询词对应的新闻信息条目；

35 计算所述新闻信息条目的第二事件属性特征；

将所述第一事件属性特征与所述第二事件属性特征进行比较，根据比较结果计算所述查询词的事件时效值；

根据所述查询词的事件时效值的大小，确定所述新闻信息条目插入到所述查询

词对应的搜索结果页上的排名位置。

6、根据权利要求 5 所述的方法，其中，所述根据所述查询词获取相关的实体词的步骤包括：

从所述查询词中提取所述实体词。

5 7、根据权利要求 5 至 6 中任一项所述的方法，其中，所述根据所述查询词获取相关的实体词的步骤包括：

根据所述查询词进行扩展得到相关的实体词。

8、根据权利要求 5 至 7 中任一项所述的方法，其中，在所述获取所述实体词的第一事件属性特征之前，还包括：

10 抓取新闻信息页面，并根据包含所述实体词的新闻信息页面计算所述第一事件属性特征。

9、根据权利要求 5 至 8 中任一项所述的方法，其中，所述计算所述新闻信息条目的第二事件属性特征的步骤包括：

根据所述新闻信息条目对应的页面，计算所述第二事件属性特征。

15 10、根据权利要求 5 至 9 中任一项所述的方法，其中，所述获取所述查询词对应的新闻信息条目的步骤包括：

从已抓取的新闻信息页面中查询出所述查询词对应的新闻信息页面，并根据所述查询词对应的新闻信息页面设置所述新闻信息条目。

20 11、根据权利要求 5 至 10 中任一项所述的方法，其中，所述获取所述查询词对应的新闻信息条目的步骤包括：

通过搜索引擎搜索所述查询词对应的多个搜索结果条目，并从所述多个搜索结果条目中提取出所述新闻信息条目。

12、一种突发时效性查询词识别系统，其包括：

25 变化趋势计算模块，用于分别计算查询词在第一段时间内和第二段时间内的查询量变化趋势，所述第二段时间为所述第一段时间的一部分；

识别模块，用于比较所述查询词在所述第一段时间和所述第二段时间内的查询量变化趋势之间的差距，并根据所述差距的大小判断所述查询词是否为突发时效性查询词。

13、根据权利要求 12 所述的突发时效性查询词识别系统，其中，还包括：

30 查询量统计模块，用于按预设长度的时间段，连续多次统计所述查询词的查询量；

所述变化趋势计算模块取最近的连续 m 个时间段为所述第一段时间，并根据所述查询词在所述 m 个时间段内的查询量，确定所述查询词在所述第一段时间内的查询量变化趋势， m 为整数；

35 所述变化趋势计算模块取最近的连续 n 个时间段为所述第二段时间，并根据所述查询词在所述 n 个时间段内的查询量，确定所述查询词在所述第二段时间内的查询量变化趋势， n 为小于 m 的整数。

14、根据权利要求 12 所述的突发时效性查询词识别系统，其中，所述识别模块取反映所述查询词在所述第一段时间内的查询量变化趋势的第一轨迹线，以及反映

所述查询词在所述第二段时间内的查询量变化趋势的第二轨迹线，并将所述第一轨迹线以及所述第二轨迹线的夹角作为所述差距。

15、根据权利要求 12 至 14 中任一项所述的突发时效性查询词识别系统，其中，还包括：

5 搜索结果条目获取模块，用于获取所述查询词对应的多个搜索结果条目，所述多个搜索结果条目中包含新闻信息条目；

权值设置模块，用于为所述多个搜索结果条目分别设置权值；

权值调整模块，用于在所述查询词为突发时效性查询词时，增大所述新闻信息条目的权值；

10 排名位置确定模块，用于根据所述多个搜索结果条目的权值大小，在所述查询词对应的搜索结果页上，确定所述多个搜索结果条目的排名位置。

16、一种在搜索结果页上插入新闻信息条目的系统，其包括：

实体词获取模块，用于接收用户的查询词，并根据所述查询词获取相关的实体词；

15 第一事件属性特征获取模块，用于获取所述实体词的第一事件属性特征；

新闻信息条目获取模块，用于获取所述查询词对应的新闻信息条目；

第二事件属性特征计算模块，用于计算所述新闻信息条目的第二事件属性特征；

事件时效值计算模块，用于将所述第一事件属性特征与所述第二事件属性特征进行比较，根据比较结果计算所述查询词的事件时效值；

20 排名位置确定模块，用于根据所述查询词的事件时效值的大小，确定所述新闻信息条目插入到所述查询词对应的搜索结果页上的排名位置。

17、根据权利要求 16 所述的系统，其中，

所述实体词获取模块从所述查询词中提取所述实体词。

18、根据权利要求 16 至 17 中任一项所述的系统，其中，

25 所述实体词获取模块根据所述查询词进行扩展得到相关的实体词。

19、根据权利要求 16 至 18 中任一项所述的系统，其中，

所述第一事件属性获取模块抓取新闻信息页面，并根据包含所述实体词的新闻信息页面，计算所述第一事件属性特征。

20、根据权利要求 16 至 19 中任一项所述的系统，其中，

30 所述第二事件属性特征计算模块根据所述新闻信息条目对应的页面，计算所述第二事件属性特征。

21、根据权利要求 16 至 20 中任一项所述的系统，其中，

所述新闻信息条目获取模块从已抓取的新闻信息页面中查询出所述查询词对应的新闻信息页面，并根据所述查询词对应的新闻信息页面设置所述新闻信息条目。

35 22、根据权利要求 16 至 21 中任一项所述的系统，其中，

所述新闻信息条目获取模块通过搜索引擎搜索所述查询词对应的多个搜索结果条目，并从所述多个搜索结果条目中提取出所述新闻信息条目。

23、一种计算机程序，包括计算机可读代码，当所述计算机可读代码在计算设备上运行时，导致所述计算设备执行根据权利要求 1 至 11 中任一项所述的方法。

24、一种计算机可读介质，其中存储了如权利要求 23 所述的计算机程序。

1/7

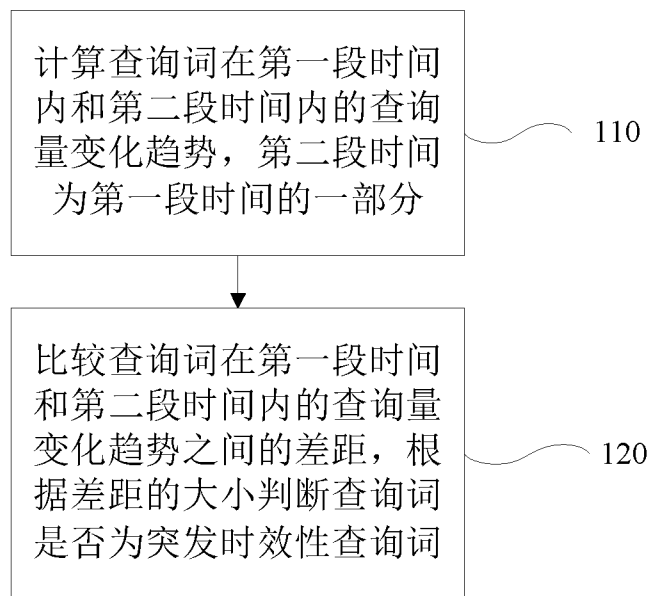


图 1

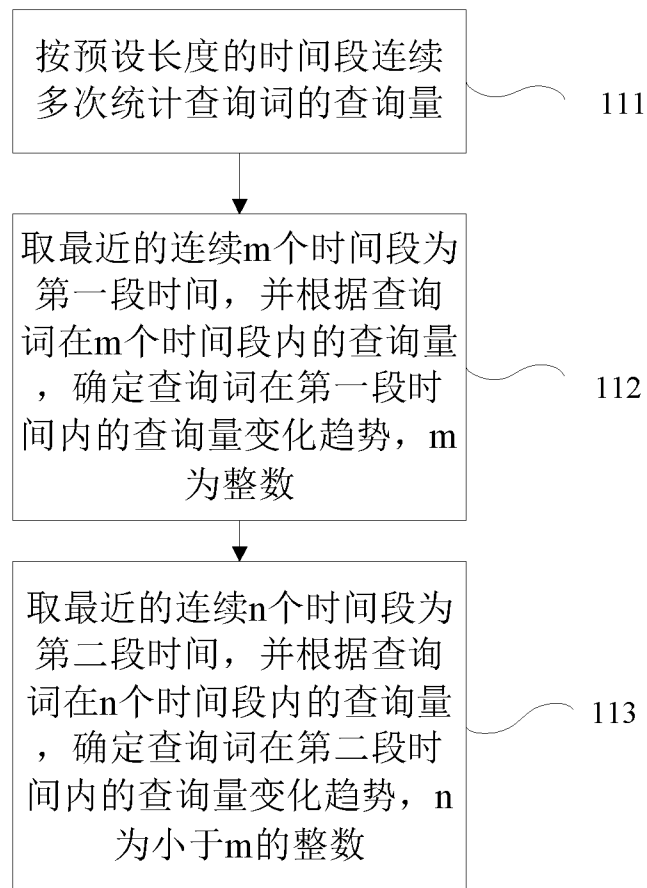


图 2

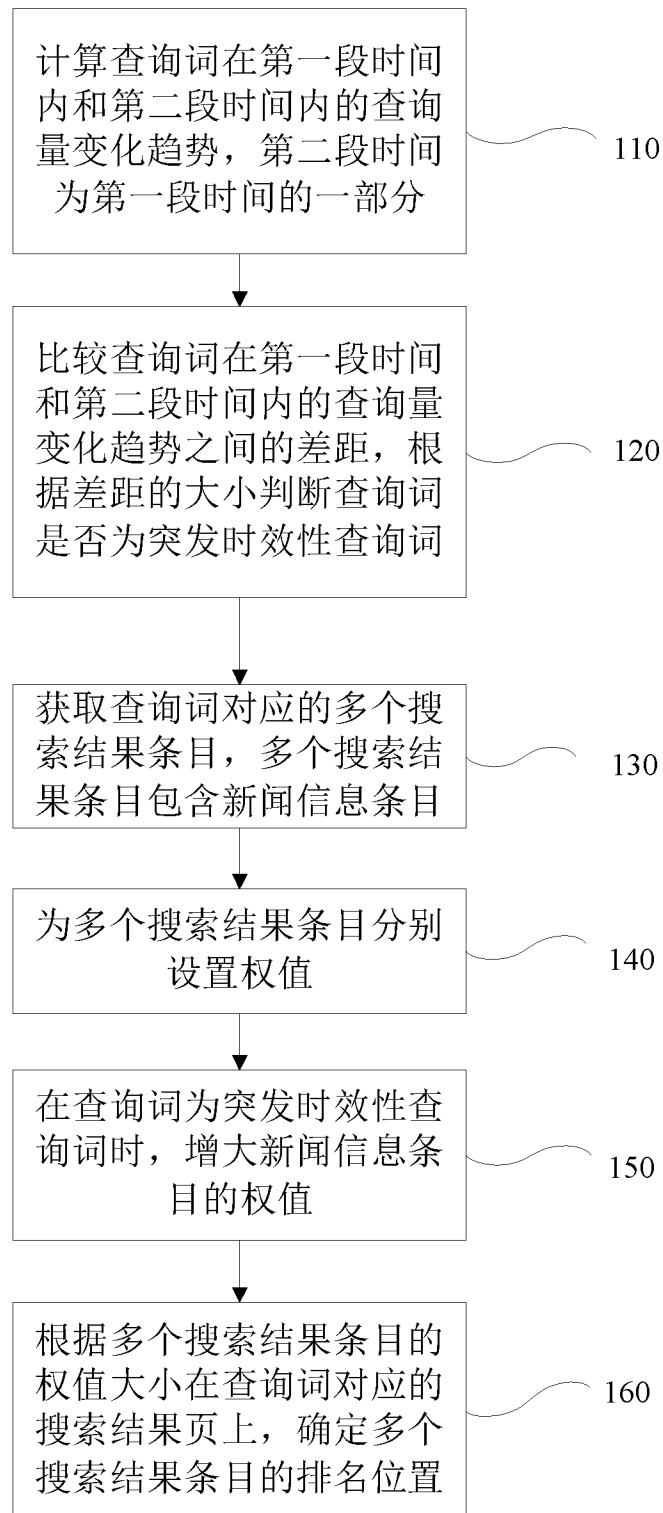


图 3

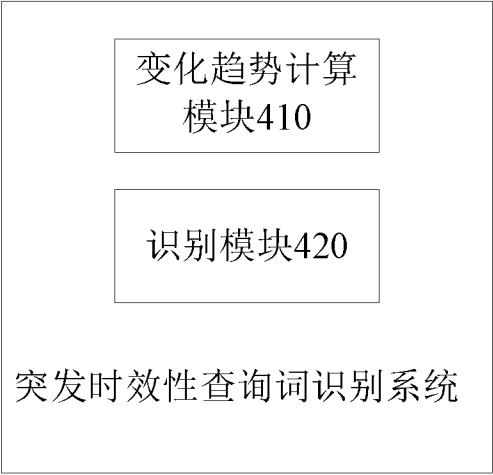


图 4



图 5



图 6

5/7

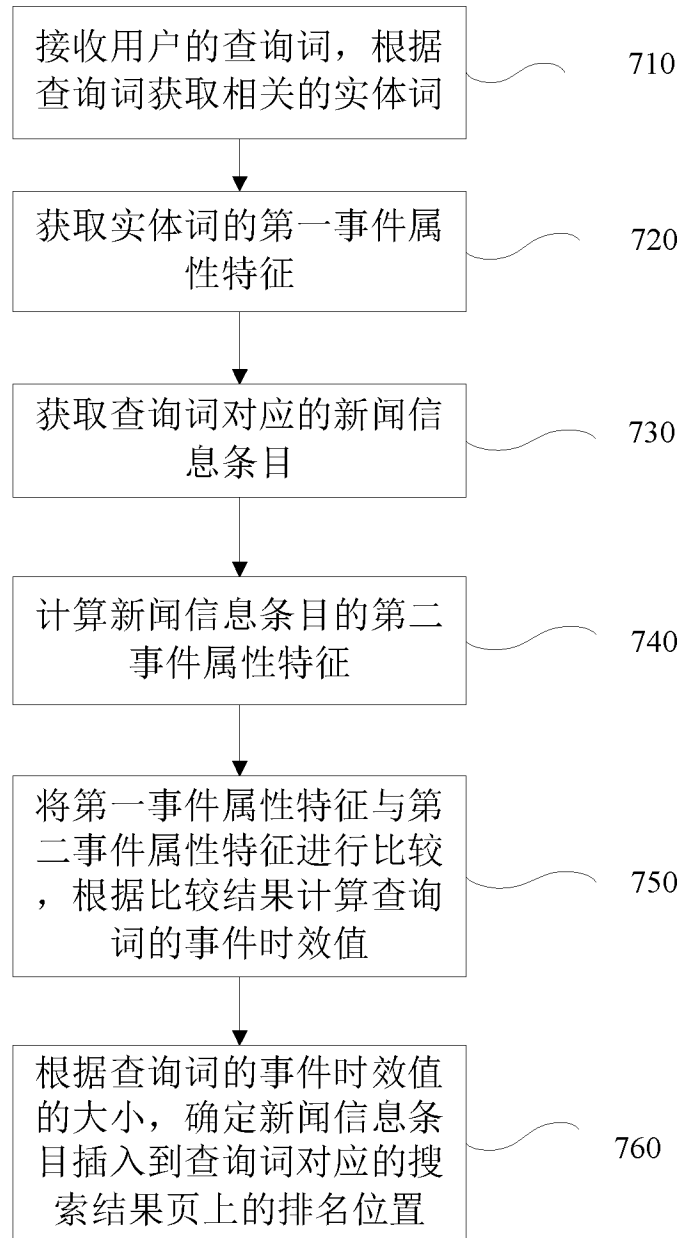


图 7

6/7



图 8

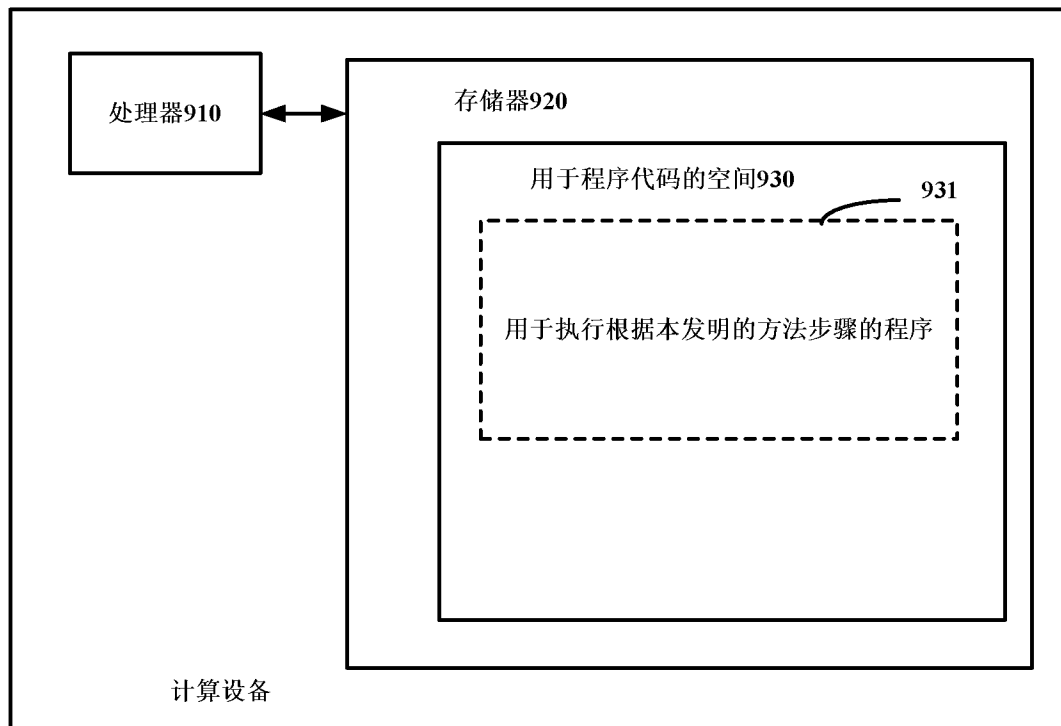


图 9

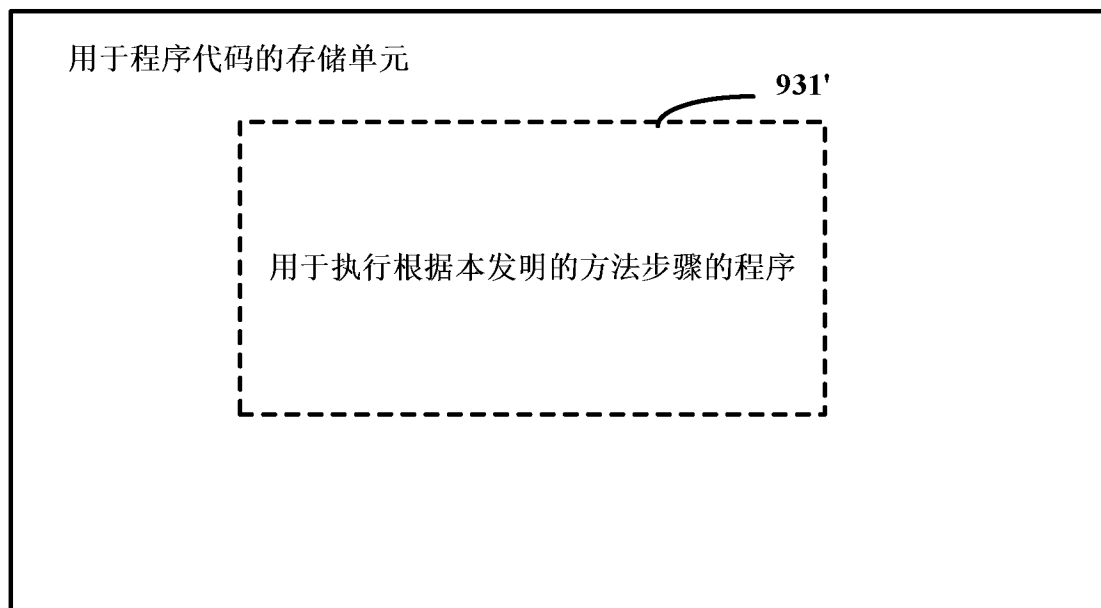


图 10

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2015/079044

A. CLASSIFICATION OF SUBJECT MATTER

G06F 17/30 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI, EPODOC, CNKI, CNPAT, IEEE: hotspot, time period, query volume, difference, entity word, key word, burst, time w line, heat+, query, word, change w trend, long w term, short w term, compar+, news, attribute, character, rank+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 103995865 A (BEIJING QIHOO TECHNOLOGY CO., LTD. et al.), 20 August 2014 (20.08.2014), claims 1-8	1-4, 12-15, 23-24
PX	CN 103984757 A (BEIJING QIHOO TECHNOLOGY CO., LTD. et al.), 13 August 2014 (13.08.2014), claims 1-10	5-11, 16-24
A	CN 103226550 A (TENCENT TECHNOLOGY (SHENZHEN) CO., LTD.), 31 July 2013 (31.07.2013), description, paragraphs [0053]-[0059] and [0070]	1-4, 12-15, 23-24
A	CN 102955829 A (BEIJING BAIDU NETWORK TECHNOLOGY CO., LTD.), 06 March 2013 (06.03.2013), description, paragraphs [0024]-[0029]	5-11, 16-24
A	CN 102004792 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.), 06 April 2011 (06.04.2011), the whole document	1-4, 12-15, 23-24
A	US 8577866 B1 (GOOGLE INC.), 05 November 2013 (05.11.2013), the whole document	5-11, 16-24

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 10 July 2015 (10.07.2015)	Date of mailing of the international search report 20 August 2015 (20.08.2015)
Name and mailing address of the ISA/CN: State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No.: (86-10) 62019451	Authorized officer MA, Chunli Telephone No.: (86-10) 82245490

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2015/079044

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

[1] The first group of claims 1-4 and 12-15 relate to a method for recognizing a burst timeliness search term; and

[2] The second group of claims 5-11 and 16-22 relate to a ranking method for inserting news information articles on a search result page;

[3] the above-mentioned two groups of claims do not share a same or corresponding special technical feature, thus they do not meet the requirement of unity.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☒ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

International application No.
PCT/CN2015/079044

Form PCT/ISA/210 (patent family annex) (July 2009)

A. 主题的分类 G06F 17/30 (2006.01) i 按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类																							
B. 检索领域 检索的最低限度文献 (标明分类系统和分类号) G06F 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用)) WPI, EPODOC, CNKI, CNPAT, IEEE: 突发, 时效, 热度, 热点, 查询, 词, 时间段, 长期, 短期, 查询量, 变化趋势, 差距, 比较, 新闻, 实体词, 关键词, 属性, 特征, 排名, 排序, burst, time w line, heat+, query, word, change w trend, long w term, short w term, compar+, news, attribute, character, rank+																							
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 103995865 A (北京奇虎科技有限公司 等) 2014年 8月 20日 (2014 - 08 - 20) 权利要求1-8</td> <td>1-4, 12-15, 23-24</td> </tr> <tr> <td>PX</td> <td>CN 103984757 A (北京奇虎科技有限公司 等) 2014年 8月 13日 (2014 - 08 - 13) 权利要求1-10</td> <td>5-11, 16-24</td> </tr> <tr> <td>A</td> <td>CN 103226550 A (腾讯科技深圳有限公司) 2013年 7月 31日 (2013 - 07 - 31) 说明书[0053]-[0059]、[0070]段</td> <td>1-4, 12-15, 23-24</td> </tr> <tr> <td>A</td> <td>CN 102955829 A (北京百度网讯科技有限公司) 2013年 3月 6日 (2013 - 03 - 06) 说明书[0024]-[0029]段</td> <td>5-11, 16-24</td> </tr> <tr> <td>A</td> <td>CN 102004792 A (百度在线网络技术北京有限公司) 2011年 4月 6日 (2011 - 04 - 06) 全文</td> <td>1-4, 12-15, 23-24</td> </tr> <tr> <td>A</td> <td>US 8577866 B1 (GOOGLE INC.) 2013年 11月 5日 (2013 - 11 - 05) 全文</td> <td>5-11, 16-24</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN 103995865 A (北京奇虎科技有限公司 等) 2014年 8月 20日 (2014 - 08 - 20) 权利要求1-8	1-4, 12-15, 23-24	PX	CN 103984757 A (北京奇虎科技有限公司 等) 2014年 8月 13日 (2014 - 08 - 13) 权利要求1-10	5-11, 16-24	A	CN 103226550 A (腾讯科技深圳有限公司) 2013年 7月 31日 (2013 - 07 - 31) 说明书[0053]-[0059]、[0070]段	1-4, 12-15, 23-24	A	CN 102955829 A (北京百度网讯科技有限公司) 2013年 3月 6日 (2013 - 03 - 06) 说明书[0024]-[0029]段	5-11, 16-24	A	CN 102004792 A (百度在线网络技术北京有限公司) 2011年 4月 6日 (2011 - 04 - 06) 全文	1-4, 12-15, 23-24	A	US 8577866 B1 (GOOGLE INC.) 2013年 11月 5日 (2013 - 11 - 05) 全文	5-11, 16-24
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																					
PX	CN 103995865 A (北京奇虎科技有限公司 等) 2014年 8月 20日 (2014 - 08 - 20) 权利要求1-8	1-4, 12-15, 23-24																					
PX	CN 103984757 A (北京奇虎科技有限公司 等) 2014年 8月 13日 (2014 - 08 - 13) 权利要求1-10	5-11, 16-24																					
A	CN 103226550 A (腾讯科技深圳有限公司) 2013年 7月 31日 (2013 - 07 - 31) 说明书[0053]-[0059]、[0070]段	1-4, 12-15, 23-24																					
A	CN 102955829 A (北京百度网讯科技有限公司) 2013年 3月 6日 (2013 - 03 - 06) 说明书[0024]-[0029]段	5-11, 16-24																					
A	CN 102004792 A (百度在线网络技术北京有限公司) 2011年 4月 6日 (2011 - 04 - 06) 全文	1-4, 12-15, 23-24																					
A	US 8577866 B1 (GOOGLE INC.) 2013年 11月 5日 (2013 - 11 - 05) 全文	5-11, 16-24																					
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																							
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件																							
国际检索实际完成的日期 2015年 7月 10日		国际检索报告邮寄日期 2015年 8月 20日																					
ISA/CN的名称和邮寄地址 中华人民共和国国家知识产权局 (ISA/CN) 北京市海淀区蓟门桥西土城路6号 100088 中国 传真号 (86-10)62019451		授权官员 马春黎 电话号码 (86-10)82245490																					

第III栏 缺乏发明单一性的意见(续第1页第3项)

本国际检索单位在该国际申请中发现多项发明，即：

- [1] 第一组权利要求1-4，12-15涉及识别突发时效性查询词的方法；
- [2] 第二组权利要求5-11，16-22涉及搜索结果页上插入新闻信息条目的排序方法；
- [3] 上述两组权利要求不具有相同或相应的特定技术特征，因此不满足发明单一性的要求。

- 1. ☐ 由于申请人按时缴纳了被要求缴纳的全部附加检索费，本国际检索报告涉及全部可作检索的权利要求。
- 2. ☒ 由于无需付出有理由要求附加费的劳动即能对全部可检索的权利要求进行检索，本单位未通知缴纳任何加费。
- 3. ☐ 由于申请人仅按时缴纳了部分被要求缴纳的附加检索费，本国际检索报告仅涉及已缴费的那些权利要求，具体地说，是权利要求：
- 4. ☐ 申请人未按时缴纳被要求缴纳的附加检索费。因此，本国际检索报告仅涉及权利要求书中首先提及的发明；包含该发明的权利要求是：

对异议的意见

- ☐ 申请人缴纳了附加检索费，同时提交了异议书，适用时，缴纳了异议费。
- ☐ 申请人缴纳了附加检索费，同时提交了异议书，但未在通知书规定的时间期限内缴纳异议费。
- ☐ 缴纳附加检索费时未提交异议书。

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2015/079044

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	103995865	A	2014年 8月 20日	无	
CN	103984757	A	2014年 8月 13日	无	
CN	103226550	A	2013年 7月 31日	无	
CN	102955829	A	2013年 3月 6日	无	
CN	102004792	A	2011年 4月 6日	无	
US	8577866	B1	2013年 11月 5日	无	