

(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
2017 年 12 月 21 日 (21.12.2017)



(10) 国际公布号
WO 2017/215243 A1

(51) 国际专利分类号:
G06F 17/30 (2006.01)

(21) 国际申请号: PCT/CN2016/113161

(22) 国际申请日: 2016 年 12 月 29 日 (29.12.2016)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201610445490.5 2016 年 6 月 17 日 (17.06.2016) CN

(71) 申请人: 广州视源电子科技股份有限公司 (GUANGZHOU SHIYUAN ELECTRONICS CO., LTD.) [CN/CN]; 中国广东省广州市黄埔区云埔四路 6 号, Guangdong 510530 (CN)。

(72) 发明人: 李贤 (LI, Xian); 中国广东省广州市黄埔区云埔四路 6 号, Guangdong 510530 (CN)。 陈振安 (CHEN, Zhenan); 中国广东省广州市黄埔区云埔四路 6 号, Guangdong 510530 (CN)。 王鹏 (WANG, Peng); 中国广东省广州市黄埔区云埔四路 6 号, Guangdong 510530 (CN)。

(74) 代理人: 广州三环专利代理有限公司 (SCIHEAD PATENT AGENT CO., LTD.); 中国广东省广州市越秀区先烈中路 80 号汇华商贸大厦 1508 室, Guangdong 510070 (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, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB,

(54) Title: METHOD AND DEVICE FOR SEARCHING RESUMES

(54) 发明名称: 简历搜索方法及装置

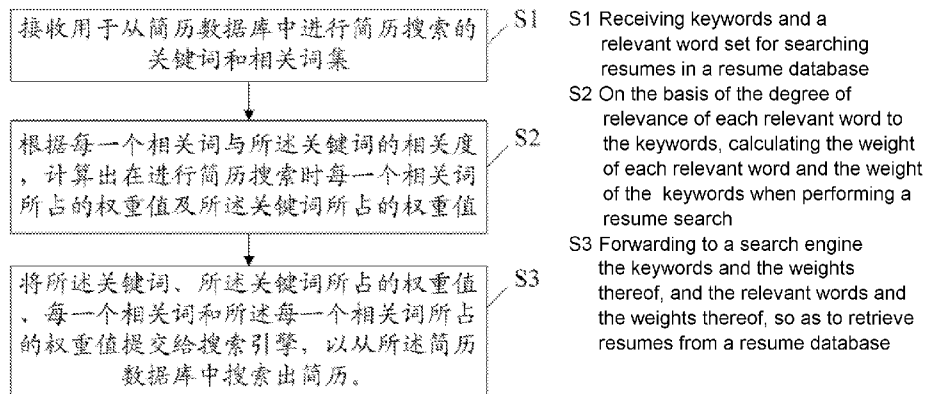


图 1

(57) Abstract: A method for searching resumes, calculating the weights of keywords and relevant words during the search process, and enhancing the precision of resume searches. The method comprises: receiving keywords and a relevant word set for searching resumes in a resume database (S1), the relevant word set comprising a number of relevant words as well as the degree of relevance of each relevant word in respect of the keywords; on the basis of the degree of relevance of each relevant word to the keywords, calculating the weight of each relevant word and the weight of the keywords when performing a resume search (S2); forwarding to a search engine the keywords and the weights thereof, and the relevant words and the weights thereof, so as to retrieve resumes from a resume database (S3).

(57) 摘要: 一种简历搜索方法和装置, 计算出关键词和相关词在搜索过程的权重, 提高检索简历的准确度; 该方法包括: 接收用于从简历数据库中进行简历搜索的关键词和相关词集 (S1); 其中, 所述相关词集记载有多个相关词, 以及每一个相关词与所述关键词的相关度; 根据每一个相关词与所述关键词的相关度, 计算出在进行简历搜索时每一个相关词所占的权重值及所述关键词所占的权重值 (S2); 将所述关键词、所述关键词所占的权重值、每一个相关词和所述每一个相关词所占的权重值提交给搜索引擎, 以从所述简历数据库中搜索出简历 (S3)。

GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KH, KN, KP, KR, KW, 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, 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。

(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, 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)。

本国际公布:

— 包括国际检索报告 (条约第21条(3))。

说明书

简历搜索方法及装置

技术领域

本发明涉及计算机信息检索领域，尤其涉及一种简历搜索方法及装置。

背景技术

在以往对简历的搜索匹配过程中，存在两种方法：一种是直接关键词检索；另一种是关键词+相关词检索；但是发明人在实施过程发现上述方案存在以下问题：

对于第一种方案，仅考虑关键词的检索过程中出现的词频来获取检索结果，即，偏向与该关键词相关的技能的简历排序在前，难以综合考虑简历的综合技能，不便于综合评估简历本人的综合技能。

对于第二种方案，在检索过程中关键词和相关词命中的频次决定了搜索出来的简历的排序，则检索结果极有可能偏向某个频次较高的相关词，另一方面，当相关词的数量大大多于关键词时，则在检索过程中关键词可能会被相关词淹没，难以突出关键词相对于相关词在搜索出简历过程中作为主要作检索关键词的作用。

发明内容

本发明实施例提出一种简历搜索方法，计算出关键词和相关词在搜索过程的权重，提高检索简历的准确度。

本发明实施例提供一种简历搜索方法，包括：

接收用于从简历数据库中进行简历搜索的关键词和相关词集；其中，所述相关词集记载有多个相关词，以及每一个相关词与所述关键词的相关度；

根据每一个相关词与所述关键词的相关度，计算出在进行简历搜索时每一个相关词所占的权重值及所述关键词所占的权重值；

将所述关键词、所述关键词所占的权重值、所述每一个相关词和所述每一个相关词所占的权重值提交给搜索引擎，以从所述简历数据库中搜索出简历。

进一步地，所述根据每一个相关词与所述关键词的相关度，计算出在进行简历搜索时每一个相关词所占的权重值及所述关键词所占的权重值，具体为：

根据基准分值公式 $S_i = \frac{r_i - r_{\min} + 1}{r_{\max} - r_{\min} + 1}$, 计算出所述相关词集中第 i 个相关词的基准分值 S_i ;

其中, r_i 为所述相关词集中的第 i 个相关词的相关度; $r_{\min}$ 为所述相关词集所记载的所有相关词的相关度的最小值; $r_{\max}$ 为所述所有相关词的相关度的最大值;

对所述所有相关词的基准分值进行求和, 获得基准分值总值 sum ;

根据第一权重公式 $W_A = 5 + \log_{1.5}(sum + 1)$, 计算出所述关键词在进行简历搜索时所占的权重值 W_A ; 其中, A 为所述关键词;

判断所述第 i 个相关词的基准分值是否大于 W_A/n ; 若是, 则根据第二权重公式

$W_i = \frac{W_A}{n \times \log_{10}(k + 10)}$, 计算出所述第 i 个相关词在进行简历搜索时所占的权重值 W_i ; 若否,

则根据第三权重公式 $W_i = \frac{S_i}{\log_{10}(k + 10)}$, 计算出所述第 i 个相关词在进行简历搜索时所占的权重值 W_i ; 其中, n 为权重系数, k 为所述相关词集中的相关词的数量。

再进一步地, 在所述判断所述第 i 个相关词的基准分值是否大于 W_A/n 之前, 还包括:

根据均值公式 $\bar{S} = \frac{\bar{r}}{r_{\max} - r_{\min} + 1}$, 计算基准分值均值 $\bar{S}$; 其中, $\bar{r}$ 为所述所有相关词的相关度的平均值;

对于所述相关词集中的第 i 个相关词, 判断所述第 i 个相关词的基准分值 S_i 是否大于所述

基准分值均值 $\bar{S}$;

若是, 通过更新公式 $S_i = 10 \times (S_i - \bar{S})$, 更新所述第 i 个相关词的基准分值 S_i 。

更进一步地, 所述将所述关键词、所述关键词所占的权重值、每一个相关词和所述每一个相关词所占的权重值提交给搜索引擎, 以从所述简历数据库中搜索出简历, 具体为:

根据搜索引擎的权重关联格式, 将所述关键词所占的权重值和所述关键词相关联作为第一组合, 以及将所述每一个相关词所占的权重值与其对应的相关词相关联作为第二组合;

将所述第一组合和所述第二组合提交给搜索引擎, 以供所述搜索引擎从所述简历数据库中搜索出简历, 并根据所述搜索引擎内置的排序算法显示搜索出的简历; 其中, 所述简历中包含的所述关键词和所述每一个相关词以高亮形式显示。

相应地, 本发明实施例还提供一种简历搜索装置, 包括:

接收模块, 用于接收用于从简历数据库中进行简历搜索的关键词和相关词集; 其中, 所

述相关词集记载有多个相关词，以及每一个相关词与所述关键词的相关度；

权重计算模块，用于根据每一个相关词与所述关键词的相关度，计算出在进行简历搜索时每一个相关词所占的权重值及所述关键词所占的权重值；

搜索模块，用于将所述关键词、所述关键词所占的权重值、所述每一个相关词和所述每一个相关词所占的权重值提交给搜索引擎，以从所述简历数据库中搜索出简历。

进一步地，所述权重计算模块具体包括：

基准分值计算单元，用于根据基准分值公式 $S_i = \frac{r_i - r_{\min} + 1}{r_{\max} - r_{\min} + 1}$ ，计算出所述相关词集中第 i 个相关词的基准分值 S_i ；其中， r_i 为所述相关词集中的第 i 个相关词的相关度； $r_{\min}$ 为在所述相关词集中所记载的所有相关词的相关度的最小值； $r_{\max}$ 为在所述所有相关词的相关度的最大值；

求和计算单元，用于对所述所有相关词的基准分值进行求和，获得基准分值总值 sum ；

关键词权重计算单元，用于根据第一权重公式 $W_A = 5 + \log_{1.5}(sum + 1)$ ，计算出所述关键词在进行简历搜索时所占的权重值 W_A ；其中， A 为所述关键词；

相关词权重计算单元，用于判断所述第 i 个相关词的基准分值是否大于 $\frac{W_A}{n}$ ；若是，则根据第二权重公式 $W_i = \frac{W_A}{n \times \log_{10}(k + 10)}$ ，计算出所述第 i 个相关词在进行简历搜索时所占的权重值 W_i ；若否，则根据第三权重公式 $W_i = \frac{S_i}{\log_{10}(k + 10)}$ ，计算出所述第 i 个相关词在进行简历搜索时所占的权重值 W_i ；其中， n 为权重系数； k 为所述相关词集中的相关词的数量。

再进一步地，所述权重计算模块还包括：

均值计算单元，用于在所述判断所述第 i 个相关词的基准分值是否大于 $\frac{W_A}{n}$ 之前，根据均值公式 $\bar{S} = \frac{\bar{r}}{r_{\max} - r_{\min} + 1}$ ，计算基准分值均值 $\bar{S}$ ；其中， $\bar{r}$ 为所述所有相关词的相关度的平均值；

判断单元，用于对于所述相关词集中的第 i 个相关词，判断所述第 i 个相关词的基准分值 S_i 是否大于所述基准分值均值 $\bar{S}$ ；

更新单元，用于当判断所述第 i 个相关词的基准分值 S_i 大于所述基准分值均值 $\bar{S}$ 时，通过更新公式 $S_i = 10 \times (S_i - \bar{S})$ ，更新所述第 i 个相关词的基准分值 S_i 。

更进一步地，所述搜索模块具体包括：

权重相联单元，用于根据搜索引擎的权重关联格式，将所述关键词所占的权重值和所述关键词相关联作为第一组合，以及将所述每一个相关词所占的权重值与其对应的相关词相关联作为第二组合；

搜索显示单元，用于将所述第一组合和所述第二组合提交给搜索引擎，以供所述搜索引擎从所述简历数据库中搜索出简历，并根据所述搜索引擎内置的排序算法显示搜索出的简历；其中，所述简历中包含的所述关键词和所述每一个相关词以高亮形式显示。

实施本发明实施例，具有如下有益效果：

本发明实施例提供的简历搜索方法和装置，通过划分关键词和相关词在简历搜索过程中所占的权重值，使得既能通过相关词检索到的简历，同时也能泛化关键词，突出关键词在检索过程中的主体地位，从而在检索过程中可以更准确地找到与关键词对应领域下匹配的简历，使得检索出来的简历更为合理。

附图说明

图 1 是本发明提供的简历搜索方法的一个实施例的流程示意图；

图 2 是本发明提供的简历搜索装置的一个实施例的结构示意图；

图 3 是本发明提供的简历搜索装置的权重计算模块的一个实施例的结构示意图；

图 4 是本发明提供的简历搜索装置的搜索模块的一个实施例的结构示意图。

具体实施方式

下面将结合本发明实施例中的附图，对本发明实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

参见图 1，是本发明提供的简历搜索方法的一个实施例的流程示意图，该方法包括步骤 S1 至 S3，具体如下：

S1，接收用于从简历数据库中进行简历搜索的关键词和相关词集；其中，所述相关词集记载有多个相关词，以及每一个相关词与所述关键词的相关度；

S2，根据每一个相关词与所述关键词的相关度，计算出在进行简历搜索时每一个相关词所占的权重值及所述关键词所占的权重值；

S3，将所述关键词、所述关键词所占的权重值、所述每一个相关词和所述每一个相关词所占的权重值提交给搜索引擎，以从所述简历数据库中搜索出简历。

在具体实施时, 上述步骤 S2 的实施方式如下:

根据基准分值公式 $S_i = \frac{r_i - r_{\min} + 1}{r_{\max} - r_{\min} + 1}$, 计算出所述相关词集中第 i 个相关词的基准分值 S_i ;

其中, r_i 为所述相关词集中的第 i 个相关词的相关度; $r_{\min}$ 为在所述相关词集中所记载的所有相关词的相关度的最小值; $r_{\max}$ 为在所述所有相关词的相关度的最大值;

对所述所有相关词的基准分值进行求和, 获得基准分值总值 sum ;

根据第一权重公式 $W_A = 5 + \log_{1.5}(sum + 1)$, 计算出所述关键词在进行简历搜索时所占的权重值 W_A ; 其中, A 为所述关键词;

判断所述第 i 个相关词的基准分值是否大于 W_A/n ; 若是, 则根据第二权重公式

$W_i = \frac{W_A}{n \times \log_{10}(k + 10)}$, 计算出所述第 i 个相关词在进行简历搜索时所占的权重值 W_i ; 若否,

则根据第三权重公式 $W_i = \frac{S_i}{\log_{10}(k + 10)}$, 计算出所述第 i 个相关词在进行简历搜索时所占的权重值 W_i ; 其中, n 为权重系数, k 为所述相关词集中的相关词的数量。

需要说明的是, 计算相关词集中每个相关词的基准分值的目的是将相关词的相关度基于同一个基准, 以一个分值的形式, 表示该相关词与关键词的相关程度, 该基准值优选为 $r_{\max} - r_{\min} + 1$; 第一权重公式采用对数形式的公式, 在基准分值总值增长过快时, 关键词的权重值仍保持缓和增长, 因而, 可以在相关词相对较多的情况下, 避免关键词所占权重过大这种情况的出现; 在进行相关词的权重计算时, 分别采用第二权重公式和第三权重公式进行计算, 是考虑到当相关词的基准分值过大和过小这两种情况出现时, 避免计算出该相关词的权重值大于关键词的权重值, 或者是大大小于关键词的权重值, 也就是说在检索简历过程中, 能避免相关词过多把关键词淹没、以及出现关键词过于突出的情况; 权重系数的设置, 来源于该相关词是从多少个词库中获取相关词的而决定, 例如, 从一个词库中获取相关词, 则 n 的取值为 2; 分别从两个词库中获取相关词, 则 n 的取值为 4。

再进一步地, 在上述步骤 S2 的实施方式中的在所述判断所述第 i 个相关词的基准分值是否大于 W_A/n 之前, 还包括:

根据均值公式 $\bar{S} = \frac{\bar{r}}{r_{\max} - r_{\min} + 1}$, 计算基准分值均值 $\bar{S}$; 其中, $\bar{r}$ 为所述所有相关词的相关度的平均值;

对于所述相关词集中的第 i 个相关词, 判断所述第 i 个相关词的基准分值 S_i 是否大于所述

基准分值均值 $\bar{S}$;

若是, 通过更新公式 $S_i = 10 \times (S_i - \bar{S})$, 更新所述第 i 个相关词的基准分值 S_i 。

需要说明的是, 将该相关词集中的相关词的相关度均值基于同一个基准, 且该基准与上述基准相同, 目的是以一个分值的形式, 表示该相关词集与关键词的平均相关程度; 当一个相关词与关键词的相关程度 (即基准分值) 大于该相关词集与关键词的平均相关程度 (即, 基准分值均值) 时, 可将该相关词与关键词的相关程度, 与该相关词集与关键词的平均相关程度两个相关程度的差异值放大 10 倍作为该相关词与关键词的相关程度的分值; 反之, 则保持该相关词与关键词的相关程度的原分值, 这种方式更能确切地表达该相关词与关键词的相关程度。

更进一步地, 上述步骤 S3 的具体实施方式为:

根据搜索引擎的权重关联格式, 将所述关键词所占的权重值和所述关键词相关联作为第一组合, 以及将所述每一个相关词所占的权重值与其对应的相关词相关联作为第二组合;

将所述第一组合和所述第二组合提交给搜索引擎, 以供所述搜索引擎从所述简历数据库中搜索出简历, 并根据所述搜索引擎内置的排序算法显示搜索出的简历; 其中, 所述简历中包含的所述关键词和所述每一个相关词以高亮形式显示。

需要说明的是, 一般采用的搜索引擎为 solr 搜索引擎, 则上述权重关联格式, 可依据如下格式进行关联: 关键词^关键词权重、相关词 1^相关词 1 权重、相关词 2^相关词 2 权重.....

实施本发明实施例的简历搜索方法, 通过调整相关词与关键词的相关程度, 使相关词与关键词之间的关系更为合理, 并依据表示该相关程度的基准分值计算出关键词和相关词在简历搜索过程中所占的权重值, 使得既能通过相关词检索到的简历, 同时也能泛化关键词, 突出关键词在检索过程中的主体地位, 从而在检索过程中可以更准确地找到与关键词对应领域下匹配的简历, 使得检索出来的简历更为合理。

相应地, 参见图 2, 是本发明提供的简历搜索装置的一个实施例的结构示意图, 该简历搜索装置, 包括:

接收模块 10, 用于接收用于从简历数据库中进行简历搜索的关键词和相关词集; 其中, 所述相关词集记载有多个相关词, 以及每一个相关词与所述关键词的相关度;

权重计算模块 20, 用于根据每一个相关词与所述关键词的相关度, 计算出在进行简历搜索时每一个相关词所占的权重值及所述关键词所占的权重值;

搜索模块 30, 用于将所述关键词、所述关键词所占的权重值、所述每一个相关词和所述每一个相关词所占的权重值提交给搜索引擎, 以从所述简历数据库中搜索出简历。

进一步地,如图3所示,图3是本发明提供的简历搜索装置的权重计算模块的一个实施例的结构示意图,所述权重计算模块20具体包括:

基准分值计算单元21,用于根据基准分值公式 $S_i = \frac{r_i - r_{\min} + 1}{r_{\max} - r_{\min} + 1}$, 计算出所述相关词集中第 i 个相关词的基准分值 S_i ; 其中, r_i 为所述相关词集中的第 i 个相关词的相关度; $r_{\min}$ 为在所述相关词集中所记载的所有相关词的相关度的最小值; $r_{\max}$ 为在所述所有相关词的相关度的最大值;

求和计算单元22,用于对所述所有相关词的基准分值进行求和,获得基准分值总值 sum ;

关键词权重计算单元23,用于根据第一权重公式 $W_A = 5 + \log_{1.5}(sum + 1)$, 计算出所述关键词在进行简历搜索时所占的权重值 W_A ; 其中, A 为所述关键词;

相关词权重计算单元24,用于判断所述第 i 个相关词的基准分值是否大于 W_A/n ; 若是,则根据第二权重公式 $W_i = \frac{W_A}{n \times \log_{10}(k + 10)}$, 计算出所述第 i 个相关词在进行简历搜索时所占的权重值 W_i ; 若否,则根据第三权重公式 $W_i = \frac{S_i}{\log_{10}(k + 10)}$, 计算出所述第 i 个相关词在进行简历搜索时所占的权重值 W_i ; 其中, n 为权重系数; k 为所述相关词集中的相关词的数量。

再进一步地,所述权重计算模块还包括:

均值计算单元25,用于在所述判断所述第 i 个相关词的基准分值是否大于 W_A/n 之前,根据均值公式 $\bar{S} = \frac{\bar{r}}{r_{\max} - r_{\min} + 1}$, 计算基准分值均值 $\bar{S}$; 其中, $\bar{r}$ 为所述所有相关词的相关度的平均值;

判断单元26,用于对于所述相关词集中的第 i 个相关词,判断所述第 i 个相关词的基准分值 S_i 是否大于所述基准分值均值 $\bar{S}$;

更新单元27,用于当判断所述第 i 个相关词的基准分值 S_i 是大于所述基准分值均值 $\bar{S}$ 时,通过更新公式 $S_i = 10 \times (S_i - \bar{S})$, 更新所述第 i 个相关词的基准分值 S_i 。

更进一步地,如图4所示,图4是本发明提供的简历搜索装置的搜索模块的一个实施例的结构示意图;所述搜索模块30具体包括:

权重相联单元31,用于根据搜索引擎的权重关联格式,将所述关键词所占的权重值和所述关键词相关联作为第一组合,以及将所述每一个相关词所占的权重值与其对应的相关词相关联作为第二组合;

搜索显示单元 32，用于将所述第一组合和所述第二组合提交给搜索引擎，以供所述搜索引擎从所述简历数据库中搜索出简历，并根据所述搜索引擎内置的排序算法显示搜索出的简历；其中，所述简历中包含的所述关键词和所述每一个相关词以高亮形式显示。

实施本发明实施例的简历搜索装置，通过调整相关词与关键词的相关程度，使相关词与关键词之间的关系更为合理，并依据表示该相关程度的基准分值计算出关键词和相关词在简历搜索过程中所占的权重值，使得既能通过相关词检索到的简历，同时也能泛化关键词，突出关键词在检索过程中的主体地位，从而在检索过程中可以更准确地找到与关键词对应领域下匹配的简历，使得检索出来的简历更为合理。

本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程，是可以通过计算机程序来指令相关的硬件来完成，所述的程序可存储于一计算机可读取存储介质中，该程序在执行时，可包括如上述各方法的实施例的流程。其中，所述的存储介质可为磁碟、光盘、只读存储记忆体（Read-Only Memory，ROM）或随机存储记忆体（Random Access Memory，RAM）等。

以上所述是本发明的优选实施方式，应当指出，对于本技术领域的普通技术人员来说，在不脱离本发明原理的前提下，还可以做出若干改进和润饰，这些改进和润饰也视为本发明的保护范围。

权利要求书

1、一种简历搜索方法，其特征在于，包括：

接收用于从简历数据库中进行简历搜索的关键词和相关词集；其中，所述相关词集记载有多个相关词，以及每一个相关词与所述关键词的相关度；

根据每一个相关词与所述关键词的相关度，计算出在进行简历搜索时每一个相关词所占的权重值及所述关键词所占的权重值；

将所述关键词、所述关键词所占的权重值、所述每一个相关词和所述每一个相关词所占的权重值提交给搜索引擎，以从所述简历数据库中搜索出简历。

2、如权利要求 1 所述简历搜索方法，其特征在于，所述根据每一个相关词与所述关键词的相关度，计算出在进行简历搜索时每一个相关词所占的权重值及所述关键词所占的权重值，具体为：

根据基准分值公式 $S_i = \frac{r_i - r_{\min} + 1}{r_{\max} - r_{\min} + 1}$ ，计算出所述相关词集中第 i 个相关词的基准分值 S_i ；

其中， r_i 为所述相关词集中的第 i 个相关词的相关度； $r_{\min}$ 为所述相关词集所记载的所有相关词的相关度的最小值； $r_{\max}$ 为所述所有相关词的相关度的最大值；

对所述所有相关词的基准分值进行求和，获得基准分值总值 sum ；

根据第一权重公式 $W_A = 5 + \log_{1.5}(sum + 1)$ ，计算出所述关键词在进行简历搜索时所占的权重值 W_A ；其中， A 为所述关键词；

判断所述第 i 个相关词的基准分值是否大于 $\frac{W_A}{n}$ ；若是，则根据第二权重公式

$W_i = \frac{W_A}{n \times \log_{10}(k + 10)}$ ，计算出所述第 i 个相关词在进行简历搜索时所占的权重值 W_i ；若否，

则根据第三权重公式 $W_i = \frac{S_i}{\log_{10}(k + 10)}$ ，计算出所述第 i 个相关词在进行简历搜索时所占的权重值 W_i ；其中， n 为权重系数， k 为所述相关词集中的相关词的数量。

3、如权利要求 2 所述的简历搜索方法，其特征在于，在所述判断所述第 i 个相关词的基准分值是否大于 $\frac{W_A}{n}$ 之前，还包括：

根据均值公式 $\bar{S} = \frac{\bar{r}}{r_{\max} - r_{\min} + 1}$, 计算基准分值均值 $\bar{S}$; 其中, $\bar{r}$ 为所述所有相关词的相关度的平均值;

对于所述相关词集中的第 i 个相关词, 判断所述第 i 个相关词的基准分值 S_i 是否大于所述基准分值均值 $\bar{S}$;

若是, 通过更新公式 $S_i = 10 \times (S_i - \bar{S})$, 更新所述第 i 个相关词的基准分值 S_i 。

4、如权利要求 1 所述的简历搜索方法, 其特征在于, 所述将所述关键词、所述关键词所占的权重值、每一个相关词和所述每一个相关词所占的权重值提交给搜索引擎, 以从所述简历数据库中搜索出简历, 具体为:

根据搜索引擎的权重关联格式, 将所述关键词所占的权重值和所述关键词相关联作为第一组合, 以及将所述每一个相关词所占的权重值与其对应的相关词相关联作为第二组合;

将所述第一组合和所述第二组合提交给搜索引擎, 以供所述搜索引擎从所述简历数据库中搜索出简历, 并根据所述搜索引擎内置的排序算法显示搜索出的简历; 其中, 所述简历中包含的所述关键词和所述每一个相关词以高亮形式显示。

5、一种简历搜索装置, 其特征在于, 包括:

接收模块, 用于接收用于从简历数据库中进行简历搜索的关键词和相关词集; 其中, 所述相关词集记载有多个相关词, 以及每一个相关词与所述关键词的相关度;

权重计算模块, 用于根据每一个相关词与所述关键词的相关度, 计算出在进行简历搜索时每一个相关词所占的权重值及所述关键词所占的权重值;

搜索模块, 用于将所述关键词、所述关键词所占的权重值、所述每一个相关词和所述每一个相关词所占的权重值提交给搜索引擎, 以从所述简历数据库中搜索出简历。

6、如权利要求 5 所述简历搜索装置, 其特征在于, 所述权重计算模块具体包括:

基准分值计算单元, 用于根据基准分值公式 $S_i = \frac{r_i - r_{\min} + 1}{r_{\max} - r_{\min} + 1}$, 计算出所述相关词集中第

i 个相关词的基准分值 S_i ; 其中, r_i 为所述相关词集中的第 i 个相关词的相关度; $r_{\min}$ 为在所述相关词集中所记载的所有相关词的相关度的最小值; $r_{\max}$ 为在所述所有相关词的相关度的最大值;

求和计算单元, 用于对所述所有相关词的基准分值进行求和, 获得基准分值总值 sum ;

关键词权重计算单元, 用于根据第一权重公式 $W_A = 5 + \log_{1.5}(sum+1)$, 计算出所述关键词在进行简历搜索时所占的权重值 W_A ; 其中, A 为所述关键词;

相关词权重计算单元, 用于判断所述第 i 个相关词的基准分值是否大于 W_A/n ; 若是, 则根据第二权重公式 $W_i = \frac{W_A}{n \times \log_{10}(k+10)}$, 计算出所述第 i 个相关词在进行简历搜索时所占的权重值 W_i ; 若否, 则根据第三权重公式 $W_i = \frac{S_i}{\log_{10}(k+10)}$, 计算出所述第 i 个相关词在进行简历搜索时所占的权重值 W_i ; 其中, n 为权重系数; k 为所述相关词集中的相关词的数量。

7、如权利要求 6 所述的简历搜索装置, 其特征在于, 所述权重计算模块还包括:

均值计算单元, 用于在所述判断所述第 i 个相关词的基准分值是否大于 W_A/n 之前, 根据均值公式 $\bar{S} = \frac{\bar{r}}{r_{\max} - r_{\min} + 1}$, 计算基准分值均值 $\bar{S}$; 其中, $\bar{r}$ 为所述所有相关词的相关度的平均值;

判断单元, 用于对于所述相关词集中的第 i 个相关词, 判断所述第 i 个相关词的基准分值 S_i 是否大于所述基准分值均值 $\bar{S}$;

更新单元, 用于当判断所述第 i 个相关词的基准分值 S_i 大于所述基准分值均值 $\bar{S}$ 时, 通过更新公式 $S_i = 10 \times (S_i - \bar{S})$, 更新所述第 i 个相关词的基准分值 S_i 。

8、如权利要求 5 所述的简历搜索装置, 其特征在于, 所述搜索模块具体包括:

权重相联单元, 用于根据搜索引擎的权重关联格式, 将所述关键词所占的权重值和所述关键词相关联作为第一组合, 以及将所述每一个相关词所占的权重值与其对应的相关词相关联作为第二组合;

搜索显示单元, 用于将所述第一组合和所述第二组合提交给搜索引擎, 以供所述搜索引擎从所述简历数据库中搜索出简历, 并根据所述搜索引擎内置的排序算法显示搜索出的简历; 其中, 所述简历中包含的所述关键词和所述每一个相关词以高亮形式显示。

说明书附图

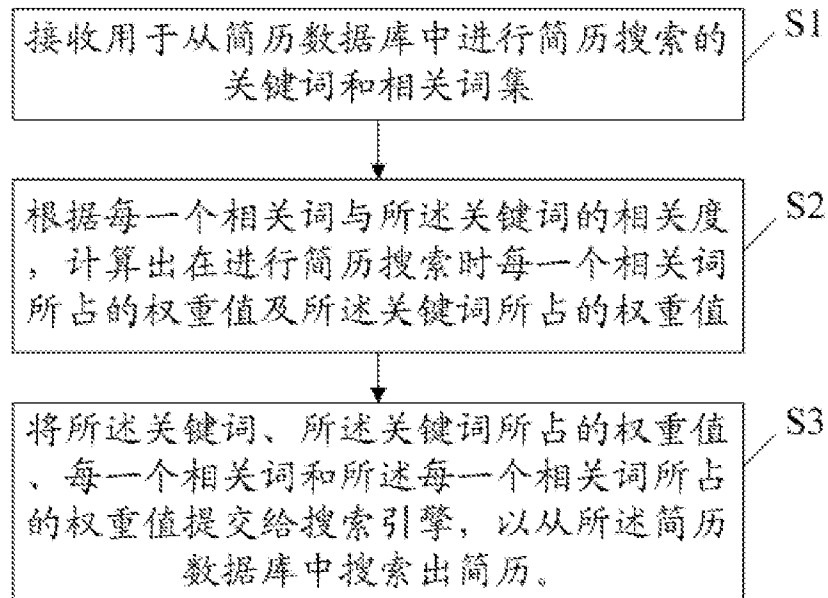


图 1

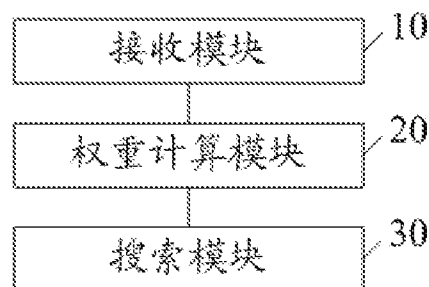


图 2

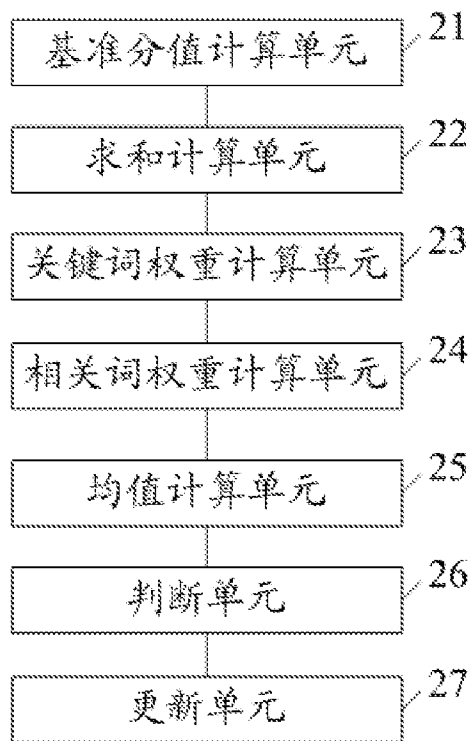


图 3

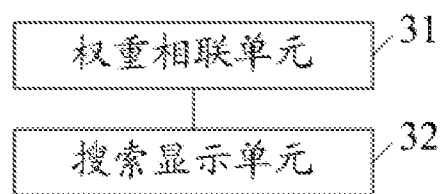


图 4

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2016/113161

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)

CNPAT, WPI, EPODOC, CNKI, IEEE, GOOGLE: query, search, key, word, term, relevant, related, weight, resume

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 106095982 A (GUANGZHOU CVTE ELECTRONIC TECHNOLOGY CO., LTD.) 09 November 2016 (09.11.2016) claims 1-8	1-8
PX	CN 105956195 A (GUANGZHOU CVTE ELECTRONIC TECHNOLOGY CO., LTD.) 21 September 2016 (21.09.2016) claims 1-10	1-8
PX	CN 106126589 A (GUANGZHOU CVTE ELECTRONIC TECHNOLOGY CO., LTD.) 16 November 2016 (16.11.2016) claims 1-10	1-8
X	CN 103425687 A (ALIBABA GROUP HOLDING LTD.) 04 December 2013 (04.12.2013) claims 1-18, and description, paragraphs [0201]-[0216]	1, 4, 5, 8
X	CN 104216931 A (KUSHENG TIANJIN SCIENCE & TECHNOLOGY CO., LTD.) 17 December 2014 (17.12.2014) description, paragraphs [0028]-[0068]	1, 4, 5, 8

☒ 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 17 February 2017	Date of mailing of the international search report 06 April 2017
Name and mailing address of the ISA 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 ZHOU, Zheng Telephone No. (86-10) 01062413664

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2016/113161

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2007034874 A (NOMURA SOGO KENKYUSHO K.K.) 08 February 2007 (08.02.2007) the whole document	1-8

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2016/113161

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 106095982 A	09 November 2016	None	
CN 105956195 A	21 September 2016	None	
CN 106126589 A	16 November 2016	None	
CN 103425687 A	04 December 2013	None	
CN 104216931 A	17 December 2014	None	
JP 2007034874 A	08 February 2007	None	

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)) CNPAT, WPI, EPDOC, CNKI, IEEE, GOOGLE: 检索, 搜索, 关键词, 相关, 关联, 扩展, 权重, 简历, query, search, key, word, term, relevant, related, weight, resume																							
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 106095982 A (广州视源电子科技有限公司) 2016年 11月 9日 (2016 - 11 - 09) 权利要求1-8</td> <td>1-8</td> </tr> <tr> <td>PX</td> <td>CN 105956195 A (广州视源电子科技有限公司) 2016年 9月 21日 (2016 - 09 - 21) 权利要求1-10</td> <td>1-8</td> </tr> <tr> <td>PX</td> <td>CN 106126589 A (广州视源电子科技有限公司) 2016年 11月 16日 (2016 - 11 - 16) 权利要求1-10</td> <td>1-8</td> </tr> <tr> <td>X</td> <td>CN 103425687 A (阿里巴巴集团控股有限公司) 2013年 12月 4日 (2013 - 12 - 04) 权利要求1-18、说明书第[0201]-[0216]段</td> <td>1、4-5、8</td> </tr> <tr> <td>X</td> <td>CN 104216931 A (酷盛天津科技有限公司) 2014年 12月 17日 (2014 - 12 - 17) 说明书第[0028]-[0068]段</td> <td>1、4-5、8</td> </tr> <tr> <td>A</td> <td>JP 2007034874 A (NOMURA SOGO KENKYUSHO KK) 2007年 2月 8日 (2007 - 02 - 08) 全文</td> <td>1-8</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 106095982 A (广州视源电子科技有限公司) 2016年 11月 9日 (2016 - 11 - 09) 权利要求1-8	1-8	PX	CN 105956195 A (广州视源电子科技有限公司) 2016年 9月 21日 (2016 - 09 - 21) 权利要求1-10	1-8	PX	CN 106126589 A (广州视源电子科技有限公司) 2016年 11月 16日 (2016 - 11 - 16) 权利要求1-10	1-8	X	CN 103425687 A (阿里巴巴集团控股有限公司) 2013年 12月 4日 (2013 - 12 - 04) 权利要求1-18、说明书第[0201]-[0216]段	1、4-5、8	X	CN 104216931 A (酷盛天津科技有限公司) 2014年 12月 17日 (2014 - 12 - 17) 说明书第[0028]-[0068]段	1、4-5、8	A	JP 2007034874 A (NOMURA SOGO KENKYUSHO KK) 2007年 2月 8日 (2007 - 02 - 08) 全文	1-8
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																					
PX	CN 106095982 A (广州视源电子科技有限公司) 2016年 11月 9日 (2016 - 11 - 09) 权利要求1-8	1-8																					
PX	CN 105956195 A (广州视源电子科技有限公司) 2016年 9月 21日 (2016 - 09 - 21) 权利要求1-10	1-8																					
PX	CN 106126589 A (广州视源电子科技有限公司) 2016年 11月 16日 (2016 - 11 - 16) 权利要求1-10	1-8																					
X	CN 103425687 A (阿里巴巴集团控股有限公司) 2013年 12月 4日 (2013 - 12 - 04) 权利要求1-18、说明书第[0201]-[0216]段	1、4-5、8																					
X	CN 104216931 A (酷盛天津科技有限公司) 2014年 12月 17日 (2014 - 12 - 17) 说明书第[0028]-[0068]段	1、4-5、8																					
A	JP 2007034874 A (NOMURA SOGO KENKYUSHO KK) 2007年 2月 8日 (2007 - 02 - 08) 全文	1-8																					
☐ 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 2017年 2月 17日		国际检索报告邮寄日期 2017年 4月 6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) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10) 62019451		授权官员 周正 电话号码 (86-10) 01062413664																					

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2016/113161

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	106095982	A	2016年 11月 9日	无	
CN	105956195	A	2016年 9月 21日	无	
CN	106126589	A	2016年 11月 16日	无	
CN	103425687	A	2013年 12月 4日	无	
CN	104216931	A	2014年 12月 17日	无	
JP	2007034874	A	2007年 2月 8日	无	