

(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
2019 年 1 月 31 日 (31.01.2019)



(10) 国际公布号
WO 2019/019674 A1

- (51) 国际专利分类号:
G06F 17/30 (2006.01)
- (21) 国际申请号: PCT/CN2018/080139
- (22) 国际申请日: 2018 年 3 月 23 日 (23.03.2018)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201710614416.6 2017年7月25日 (25.07.2017) CN
- (71) 申请人: 深圳壹账通智能科技有限公司(ONE CONNECT SMART TECHNOLOGY CO., LTD. (SHENZHEN)) [CN/CN]; 中国广东省深圳市前海深港合作区前湾一路1号A栋201室, Guangdong 518052 (CN)。
- (72) 发明人: 李武奇(LI, WUqi); 中国广东省深圳市前海深港合作区前湾一路1号A栋201室, Guangdong 518052 (CN)。
- (74) 代理人: 广州华进联合专利商标代理有限公司(ADVANCE CHINA IP LAW OFFICE); 中国广东省广州市天河区花城大道85号3901房, Guangdong 510623 (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, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, 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)。

根据细则4.17的声明:

- 关于申请人有权要求在先申请的优先权(细则4.17(iii))

(54) **Title:** WEBSITE ERROR INFORMATION PROCESSING METHOD AND APPARATUS, COMPUTER DEVICE AND STORAGE MEDIUM

(54) 发明名称: 网页错误信息处理方法、装置、计算机设备和存储介质

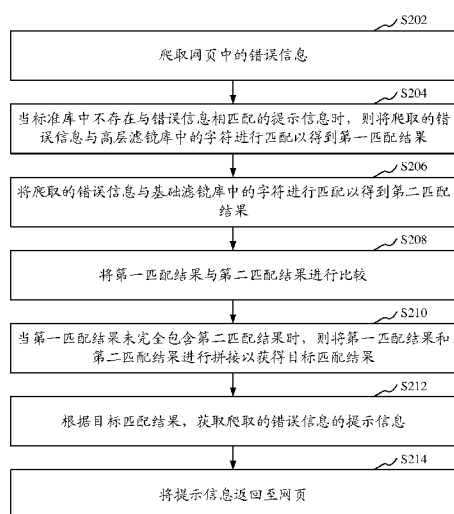


图 2

S202 CRAWL ERROR INFORMATION IN A WEBSITE
S204 WHEN PROMPT INFORMATION MATCHING THE ERROR INFORMATION DOES NOT EXIST IN A STANDARD LIBRARY, MATCH THE CRAWLED ERROR INFORMATION WITH CHARACTERS IN A HIGH-LEVEL FILTER LIBRARY TO OBTAIN A FIRST MATCHING RESULT
S206 MATCH THE CRAWLED ERROR INFORMATION WITH CHARACTERS IN A BASIC FILTER LIBRARY TO OBTAIN A SECOND MATCHING RESULT
S208 COMPARE THE FIRST MATCHING RESULT WITH THE SECOND MATCHING RESULT
S210 WHEN THE SECOND MATCHING RESULT IS NOT COMPLETELY CONTAINED IN THE FIRST MATCHING RESULT, SPLICE THE FIRST MATCHING RESULT AND THE SECOND MATCHING RESULT TO OBTAIN A TARGET MATCHING RESULT
S212 ACQUIRE, ACCORDING TO THE TARGET MATCHING RESULT, THE PROMPT INFORMATION ABOUT THE CRAWLED ERROR INFORMATION
S214 RETURN THE PROMPT INFORMATION TO THE WEBSITE

(57) **Abstract:** Disclosed is a website error information processing method. The method comprises: crawling error information in a website; when prompt information matching the error information does not exist in a standard library, matching the crawled error information with characters in a high-level filter library to obtain a first matching result, and matching same with characters in a basic filter library to obtain a second matching result; when the second matching result is not completely contained in the first matching result, splicing the first matching result and the second matching result to obtain a target matching result; acquiring, according to the target matching result, the prompt information about the crawled error information; and returning the prompt information to the website.

(57) **摘要:** 一种网页错误信息处理方法, 该方法包括: 爬取网页中的错误信息; 当标准库中不存在与错误信息相匹配的提示信息时, 则将爬取的错误信息与高层滤镜库中的字符匹配得到第一匹配结果, 与基础滤镜库中的字符匹配得到第二匹配结果; 当第一匹配结果未完全包含第二匹配结果时, 将第一匹配结果和第二匹配结果进行拼接得到目标匹配结果; 根据目标匹配结果, 获取爬取的错误信息的提示信息; 将提示信息返回至网页。

本国际公布：

- 包括国际检索报告(条约第21条(3))。

网页错误信息处理方法、装置、计算机设备和存储介质

相关申请的交叉引用

本申请要求于 2017 年 7 月 25 日提交中国专利局，申请号为 2017106144166，申请名称为“网页错误信息处理方法、装置、计算机设备和存储介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及一种网页错误信息处理方法、装置、计算机设备和存储介质。

背景技术

专用爬虫解析系统，为对要爬取的页面结构有依赖，页面结构的改变或页面的替换都会导致解析或爬取失败的爬虫解析系统。对于专用爬虫解析系统，当线上出现网站改版或新型账户类型时，发明人意识到，会出现一些新发未知页面、新发未知错误，从而会导致网络爬取系统的可用性降低且影响用户体验。传统上针对这些新发问题的处理，多采取线下修复，测试研发版本这一流程，这样耗时长。

发明内容

根据本申请公开的各种实施例，提供一种网页错误信息处理方法、装置、计算机设备和存储介质。

一种网页错误信息处理方法，包括：

爬取网页中的错误信息；

当标准库中不存在与所述错误信息相匹配的提示信息时，则将爬取的所述错误信息与高层滤镜库中的字符进行匹配以得到第一匹配结果；

将爬取的所述错误信息与基础滤镜库中的字符进行匹配以得到第二匹配结果；

将所述第一匹配结果与所述第二匹配结果进行比较；

当所述第一匹配结果未完全包含所述第二匹配结果时，则将第一匹配结果和所述第二匹配结果进行拼接以获得目标匹配结果；

根据所述目标匹配结果，获取爬取的所述错误信息的提示信息；及

将所述提示信息返回至所述网页。

一种网页错误信息处理装置，包括：

爬取模块，用于爬取网页中的错误信息；

匹配模块，用于当标准库中不存在与所述错误信息相匹配的提示信息时，则将爬取的所述错误信息与高层滤镜库中的字符进行匹配以得到第一匹配结果；将爬取的所述错误信

息与基础滤镜库中的字符进行匹配以得到第二匹配结果;

比较模块,用于将所述第一匹配结果与所述第二匹配结果进行比较;当所述第一匹配结果未完全包含所述第二匹配结果时,则将第一匹配结果和所述第二匹配结果进行拼接以获得目标匹配结果;

- 5 获取模块,用于根据所述目标匹配结果,获取爬取的所述错误信息的提示信息;及
返回模块,用于将所述提示信息返回至所述网页。

一种计算机设备,包括存储器和一个或多个处理器,所述存储器中储存有计算机可读指令,,所述计算机可读指令被一个或多个所述处理器执行时,使得所述一个或多个处理器执行以下步骤:

- 10 爬取网页中的错误信息;

当标准库中不存在与所述错误信息相匹配的提示信息时,则将爬取的所述错误信息与高层滤镜库中的字符进行匹配以得到第一匹配结果;

将爬取的所述错误信息与基础滤镜库中的字符进行匹配以得到第二匹配结果;

将所述第一匹配结果与所述第二匹配结果进行比较;

- 15 当所述第一匹配结果未完全包含所述第二匹配结果时,则将第一匹配结果和所述第二匹配结果进行拼接以获得目标匹配结果;

根据所述目标匹配结果,获取爬取的所述错误信息的提示信息;及

将所述提示信息返回至所述网页。

- 20 一个或多个存储有计算机可读指令的非易失性计算机可读指令存储介质,计算机可读指令被一个或多个处理器执行时,使得一个或多个处理器执行以下步骤:

爬取网页中的错误信息;

当标准库中不存在与所述错误信息相匹配的提示信息时,则将爬取的所述错误信息与高层滤镜库中的字符进行匹配以得到第一匹配结果;

将爬取的所述错误信息与基础滤镜库中的字符进行匹配以得到第二匹配结果;

- 25 将所述第一匹配结果与所述第二匹配结果进行比较;

当所述第一匹配结果未完全包含所述第二匹配结果时,则将第一匹配结果和所述第二匹配结果进行拼接以获得目标匹配结果;

根据所述目标匹配结果,获取爬取的所述错误信息的提示信息;及

将所述提示信息返回至所述网页。

- 30 本申请的一个或多个实施例的细节在下面的附图和描述中提出。本申请的其它特征和优点将从说明书、附图以及权利要求书变得明显。

附图说明

- 35 为了更清楚地说明本申请实施例中的技术方案,下面将对实施例中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本申请的一些实施例,对于本领域

普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其它的附图。

图 1 为根据一个或多个实施例中网页错误信息处理方法的应用场景图。

图 2 为根据一个或多个实施例中网页错误信息处理方法的流程图。

图 3 为图 2 所示实施例的步骤 S212 的流程图。

5 图 4 为根据一个或多个实施例中目标匹配结果的示意图。

图 5 为图 2 所示实施例中步骤 S204 的流程图。

图 6 为图 2 所示实施例中的步骤 S208 的流程图。

图 7 为根据一个或多个实施例中的网页错误信息处理装置的结构示意图。

图 8 为根据一个或多个实施例中的专用爬虫解析终端的结构示意图。

10

具体实施方式

为了使本申请的技术方案及优点更加清楚明白,以下结合附图及实施例,对本申请进行进一步详细说明。应当理解,此处所描述的具体实施例仅仅用于解释本申请,并不用于限定本申请。

15 请参阅图 1,图 1 为根据一个或多个实施例中网页错误信息处理方法的应用场景图,专用爬虫解析终端与数个用户终端可以进行通信,专用爬虫解析终端可以爬取数个用户终端上的错误信息,例如当用户终端处于登录界面时,用户向用户终端输入用户名和密码,当密码输入错误时,用户终端的界面会显示初始信息“密码错误”的字样,专用爬虫解析终端爬取此时用户终端的界面上显示的初始信息“密码错误”的字样,再获取与“密码错误”
20 相对应的提示信息“输入密码与正确密码不符,请重新输入”,最后专用爬虫解析终端将该提示信息“输入密码与正确密码不符,请重新输入”返回至用户终端,以进行显示。该专用爬虫解析终端以及用户终端可以但不限于各种个人计算机、笔记本电脑、智能手机、平板电脑和便携式可穿戴设备。

请参阅图 2,图 2 为根据一个或多个实施例中网页错误信息处理方法的流程图,本实施例以该方法应用到上述图 1 中的网页错误信息处理系统中的专用爬虫解析终端来举例说明。该方法具体包括如下步骤:

S202: 爬取网页中的错误信息。

具体地,错误信息是指提示用户操作存在错误的信息,例如,用户登录某一网站或客户端时,用户名输入错误所产生的“用户名输入错误”的错误信息,或者用户填写某一表格时,用户输入的数据格式错误,例如填写日期时,格式应该为“年-月-日”,但用户输入的格式为“月-日-年”,用户终端的屏幕会显示“格式错误”的错误信息。网页在此处是指
30 广义的网页,即包含浏览器界面所显示的网页,又包含客户端 APP(Application)等显示的网页。

当网页中存在错误信息时,专用爬虫解析终端可以爬取网页中显示的错误信息,以便
35 进行后续的解析工作,从而网页上可以输出统一的提示信息,使得用户可以准确地知晓需

要更改的地方，以及下一步应该如何操作。

S204: 当标准库中不存在与错误信息相匹配的提示信息时，则将爬取的错误信息与高层滤镜库中的字符进行匹配以得到第一匹配结果。

具体地，标准库是指存储有错误信息以及提示信息的匹配关系的数据库，其一般是根据现有的错误信息和提示信息整理所得到的，即错误信息和提示信息的基础库。针对不同的网页中的错误信息，均可以通过该标准库获取相应错误信息的提示信息，这样，无论用户登录何种网页，所有网页针对相同的错误所显示的提示信息均一致，以便于用户修改错误，继续后续操作，防止由于网页不同，针对同一错误出现多种提示信息，而导致用户不知如何进行下一步操作的情况发生。

高层滤镜库中的字符主要为复杂字符，例如可以为词组或短句，例如“账户异常”、“重设密码”、“控件无法显示”等，高层滤镜库的存在可以提高匹配的精度。

当标准库中存在与错误信息相匹配的提示信息时，即该错误信息已经在标准库中，且存在相应的提示信息，则专用爬虫解析终端将该提示信息发送至用户终端进行显示，从而用户可以知晓下一步该如何进行，以便于用户继续进行后续操作。

当标准库中不存在与错误信息相匹配的提示信息时，即该错误信息为新发错误，并没有相应的标准提示信息，则专用爬虫解析终端需要确定该错误信息的提示信息，即将该错误信息与高层滤镜库中的字符进行匹配，当高层滤镜库中存在与该错误信息相对应的字符时，则存储在第一匹配结果中。

S206: 将爬取的错误信息与基础滤镜库中的字符进行匹配以得到第二匹配结果。

具体地，基础滤镜库中的字符主要为简单字符，例如可以为单个词或词组，例如“网银”、“客户”、“控件”、“重设”、“异常”、“无法显示”、“页面”等，基础滤镜库的存在可以提高专用爬虫解析系统的普遍适用性。

当标准库中不存在与错误信息相匹配的提示信息时，即该错误信息为新发错误，并没有相应的标准提示信息，则专用爬虫解析终端需要确定该错误信息的提示信息，即将该错误信息与基础滤镜库中的字符进行匹配，当基础滤镜库中存在与该错误信息相对应的字符时，则存储在第二匹配结果中。可选地，错误信息可以先与高层滤镜库中的字符进行匹配，也可以先与基础滤镜库中的字符进行匹配，或者通过两个线程同时与高层滤镜库以及基础滤镜库中的字符进行匹配。

S208: 将第一匹配结果与第二匹配结果进行比较。

具体地，由于第一匹配结果是将错误信息与高层滤镜库中的字符进行匹配的结果，第二匹配结果是将错误信息与基础滤镜库中的字符进行匹配的结果，所以第一匹配结果和第二匹配结果可以存在差异，因此将第一匹配结果与第二匹配结果进行比较，以获得准确的目标匹配结果。

S210: 当第一匹配结果未完全包含第二匹配结果时，则将第一匹配结果和第二匹配结果进行拼接以获得目标匹配结果。

具体地，当第一匹配结果与第二匹配结果存在交集，但第一匹配结果与第二匹配结果又各自存在不相同的词汇时，或者当第一匹配结果与第二匹配结果不存在交集，但第一匹配结果与第二匹配结果又各自存在不相同的词汇时，则表示第一匹配结果未完全包含第二匹配结果，从而可以通过将第一匹配结果和第二匹配结果进行拼接以获得目标匹配结果，即获取第一匹配结果和第二匹配结果的合集以作为目标匹配结果，从而保证最后的目标匹配结果的完整性。

例如，例如当错误信息为“账户异常，请去柜台重新设置”时，基础滤镜库为简单字符，如“网银”，“客户”，“控件”，“重新设置”，“异常”，“无法显示”，“页面”。高层滤镜库主要为复杂字符，例如“账户异常”，“重设密码”，“控件无法显示”。通过高层滤镜库获得的第一匹配结果为“账户异常”，通过基础滤镜库获得的第二匹配结果为“异常”“重新设置”。从而第一匹配结果并未完全包含第二匹配结果，进而目标匹配结果为第一匹配结果与第二匹配结果的合集，即“账户异常”、“重新设置”。

S212: 根据目标匹配结果，获取爬取的错误信息的提示信息。

具体地，通过目标匹配结果可以得到对应于错误信息的提示信息，例如可以通过目标匹配结果首先得到对应于目标匹配结果的匹配码，然后根据匹配码来获取对应的提示信息，其中匹配码的设置是为了对提示信息进行分类，便于提示信息的查询，例如以“1”开头的匹配码表示密码错误，以“2”开头的匹配码表示验证码错误，以“4”开头的匹配码表示“账号错误”等，其中以“4”开头的匹配码“4217”表示的是“账户错误”类型下的第“217”类错误所对应的提示信息“需要重新设置账号信息”。例如上述实施例中标匹配结果为“账户异常”、“重新设置”，经查询其对应的匹配码应该为 4217，对应的提示信息为“需要重新设置账号信息”。

S214: 将提示信息返回至网页。

具体地，当获取到相应的提示信息后，可以将该提示信息发送给网页，从而网页可以显示该提示信息，以便于用户了解到下一步该如何进行操作，例如，网页显示的提示信息为“需要重新设置账号信息”时，用户则重新设置账号信息即可进行下一步的操作。

上述网页错误信息处理方法，当网页中存在错误信息时，在标准库中又没有与该错误信息相对应的提示信息时，即该错误信息为新发错误信息时，则通过高层滤镜库中的字符和基础滤镜库中的字符对该错误信息进行匹配，以确定该错误信息的目标匹配结果，从而可以得到针对该错误信息的提示信息，将该提示信息返回至网页进行显示后，用户可以根据该提示信息进行相应的修改，以继续进行网页操作，其中通过高层滤镜库来确定错误信息的目标匹配结果可以提高匹配的准确度，通过基础滤镜库来确定错误信息的目标匹配结果可以提高专用爬虫解析系统的普遍适用性，从而使得当网页中存在新发错误信息时，可以通过高层滤镜库和基础滤镜库来确定错误信息的目标匹配结果，不需要采取线下修复，测试研发版本这一流程，节约时间。

在其中一个实施例中，请参阅图 3，图 3 为图 2 所示实施例的步骤 S212 的流程图，

该步骤 S212, 即根据目标匹配结果, 获取爬取的错误信息的提示信息的步骤, 可以包括:

S302: 接收针对目标匹配结果的调整指令。

具体地, 针对目标匹配结果的调整指令是指工作人员向专用爬虫解析终端输入的调整指令, 该指令可以表示确认, 即接受专用爬虫解析终端所得到的目标匹配结果, 该指令也可以表示调整, 即针对专用爬虫解析终端所得到的目标匹配结果进行调整, 该指令还可以表示撤回, 即拒绝专用爬虫解析终端的目标匹配结果。

S304: 当调整指令表示确认时, 则根据目标匹配结果, 获取爬取的错误信息的提示信息。

具体地, 当调整指令表示确认时, 则通过目标匹配结果可以得到对应于错误信息的提示信息, 具体可以参见上文中的限定, 在此不再赘述。

S306: 当调整指令表示调整时, 则根据调整指令对目标匹配结果进行调整, 并根据调整后的目标匹配结果, 获取爬取的错误信息的提示信息。

具体地, 当调整指令表示调整时, 即专用爬虫解析终端给出的目标匹配结果存在错误, 工作人员对该目标匹配结果进行调整。参阅图 4, 图 4 为一实施例中目标匹配结果的示意图, 其中包括原始错误信息“账户异常, 请去柜台重新设置”, 匹配码“4217”, 目标匹配结果“账户异常”“重新设置”以及“确认”、“调整”、“撤回”按钮, 从而工作人员可以通过“确认”、“调整”、“撤回”按钮来输入调整指令, 例如, 当工作人员点击“调整”时, 则可以针对目标匹配结果“账户异常”“重新设置”进行调整, 例如增加“柜台”字符, 从而使得调整后的目标匹配结果为“账户异常”“柜台”“重新设置”。专用爬虫解析终端根据该调整后的目标匹配结果“账户异常”“柜台”“重新设置”获取对应的匹配码, 例如“4218”, 然后根据匹配码来获取对应的提示信息“需要到柜台重新设置账号信息”。最后专用爬虫解析终端将该提示信息“需要到柜台重新设置账号信息”返回至网页, 从而网页可以显示该提示信息, 以便于用户了解到下一步该如何进行操作, 例如, 网页显示的提示信息为“需要到柜台重新设置账号信息”时, 用户则到柜台重新设置账号信息即可进行下一步的操作。

上述实施例中, 引用工作人员的调整指令, 对专用爬虫解析终端针对新发错误自动匹配所得到的目标匹配结果进行调整, 使得目标匹配结果更加准确, 专用爬虫解析终端适应性更强。

在其中一个实施例中, 错误信息处理方法还可以包括将根据目标匹配结果获取的提示信息以及相应的错误信息关联存储至补充库的步骤。即将通过专用爬虫解析终端为新发错误所匹配的提示信息关联存储至补充库中, 当该新发错误的目标匹配结果经过调整时, 则将调整后的目标匹配结果所对应的提示信息以及相应的错误信息关联存储在补充库中。

请参阅图 5, 图 5 为图 2 所示实施例中步骤 S204 的流程图, 在该实施例中, 该步骤 S204, 即当标准库中不存在与错误信息相匹配的提示信息时, 则将爬取的错误信息与高层滤镜库中的字符进行匹配以得到第一匹配结果的步骤, 可以包括:

S502: 当标准库中不存在与错误信息相匹配的提示信息时, 则将爬取的错误信息与补充库中已存储的错误信息进行匹配。

具体地, 当标准库中不存在与错误信息相匹配的提示信息时, 即该错误信息为新发错误, 并没有相应的标准提示信息, 但是由于专用爬虫解析终端已经对一些新发错误匹配了提示信息, 并将该提示信息与新发错误存储在了补充库中, 因此为了节约时间, 专用爬虫解析终端首先根据错误信息查询补充库, 即当补充库中存在该错误信息时, 则直接根据补充库获取对应的提示信息即可, 不需要再通过高层滤镜库和基础滤镜库对该错误信息进行匹配从而获得目标匹配结果, 在根据目标匹配结果获取对应的提示信息的复杂步骤, 节约时间。

S504: 当爬取的错误信息与补充库中已存储的错误信息匹配成功时, 则获取匹配成功的已存储的错误信息相对应的提示信息。

具体地, 即该新发错误已经存储在了补充库中, 因此不再需要通过高层滤镜库和基础滤镜库对该错误信息进行匹配从而获得目标匹配结果, 在根据目标匹配结果获取对应的提示信息的复杂步骤, 而是直接获取补充库中与该爬取的错误信息相关联的提示信息即可。

例如, 上一次出现的新发错误为“账户异常, 请去柜台重新设置”, 专用爬虫解析终端匹配出的提示信息为“需要重新设置账号信息”, 经过工作人员调整后, 存储在补充库中的新发错误“账户异常, 请去柜台重新设置”的提示信息为“需要到柜台重新设置账号信息”。本次又存储新发错误“账户异常, 请去柜台重新设置”时, 则不需要通过高层滤镜库、基础滤镜库以及工作人员来进行调整, 而只需要通过查询补充库中的错误信息“账户异常, 请去柜台重新设置”的提示信息“需要到柜台重新设置账号信息”即可, 节约时间。

S506: 将提示信息返回至网页。

具体地, 当在补充库中查询到相应的提示信息后, 则直接将该提示信息返回至网页, 从而网页可以显示该提示信息, 以便于用户了解到下一步该如何进行操作, 例如, 网页显示的提示信息为“需要到柜台重新设置账号信息”时, 用户则重新设置账号信息即可进行下一步的操作。

S508: 当爬取的错误信息与补充库中已存储的错误信息匹配失败时, 则继续将错误信息与高层滤镜库中的字符进行匹配以得到第一匹配结果的步骤。

具体地, 当爬取的错误信息与补充库中已存储的错误信息匹配失败时, 即标准库和补充库中均未存储该错误信息, 即该错误信息为新发错误, 则需要通过高层滤镜库、基础滤镜库来为该错误信息匹配出相适应的提示信息。

上述实施例中, 当通过高层滤镜库和基础滤镜库为错误信息匹配出提示信息后, 则将该提示信息和对应的错误信息关联存储在补充库, 且当工作人员调整该提示信息后, 将调整后的提示信息和对应的错误信息关联存储在补充库, 这样当网页出现新发错误时, 首先根据补充库来获取与该新发错误相匹配的提示信息, 而不需要通过高层滤镜库、基础滤镜

库以及工作人员来进行调整，节约时间。

在其中一个实施例中，当第一匹配结果完全包含第二匹配结果时，则将第一匹配结果作为目标匹配结果。在其中一个实施例中，当第一匹配结果中不存在字符时，则将第二匹配结果作为目标匹配结果。

5 具体地，高层滤镜库中存储的是复杂字符，匹配出来的结果也更加精准。例如输入的原始信息为<frame name= “result” ><h2class= “aler” style= “visibility:none” >，错误提示</h2><br class= “erromsg” >用户账户异常，请重新设置密码</br><td></td></frame>。用基础滤镜库匹配出结果“账户” & “异常” & “密码” & “设置”，同时使用高层滤镜库匹配输出的结果为“账户异常” & “重新设置” & “设置密码”。基础滤镜库的匹配结果为“账户异常，设置密码”。高层滤镜库的匹配结果：“账户异常，重新设置密码”。此时，高层滤镜库的匹配结果包含基础滤镜库的匹配结果。所以高层滤镜库的匹配结果被采用。最后的目标匹配结果为“账户异常，重新设置密码”。如果基础滤镜库的匹配结果未被高级滤镜库的匹配结果全包含，则会将基础滤镜库匹配结果多出的部分和高级滤镜库匹配结果组合使用。

15 高层滤镜库虽然精度高，但是复杂词汇组容易因为个别字的变动导致匹配不成功。即第一匹配结果中不存在字符，例如当输入信息为“账户状态异常，请去柜台设置新的密码”时，高层滤镜库“账户异常”和“设置密码”都会匹配失败。此时基础滤镜库就很有使用的必要。使用基础滤镜库就可以获得输出“账户” & “异常” & “设置” & “密码”。使用基础滤镜增加了系统的普适性。

20 高层滤镜库和基础滤镜库的词语可以更改，即添加、更换、删除，以提高滤镜的适应性。例如，可以获取统计错误信息的输出的次数，当次数达到一定值时，则将该错误信息对应的提示信息增加至高层滤镜库中。

上述实施例中，第一匹配结果是通过高层滤镜库进行匹配所得到的精确结果，第二匹配结果是通过基础滤镜库进行匹配所得到的普适结果，通过判断第一匹配结果和第二匹配结果的关系，来确定最终的目标匹配结果，提高了目标匹配结果的准确率。

25 参阅图 6，图 6 为图 2 所示实施例中的步骤 S208 的流程图，该步骤 S208，即将第一匹配结果与第二匹配结果进行比较的步骤，可以包括：

S602：当第一匹配结果中存在字符时，则通过第二匹配结果中的每一字符去匹配第一匹配结果中的每一字符。

30 具体地，当第一匹配结果中存在字符时，即第一匹配记过中存在有相应的匹配结果，例如，当错误信息为“用户账户异常，请重新设置密码”时，通过高层滤镜库匹配出的第一匹配结果为“账户异常，重新设置密码”，通过基础滤镜库匹配出的第二匹配结果为“账户异常，设置密码”，则通过第二匹配结果中的每一字符去匹配第一匹配结果中的每一字符，例如第二匹配结果中的“账户”存在于第一匹配结果中，第二匹配结果中的“异常”也存在与第一匹配结果中，第二匹配结果中的“设置”也存在于第一匹配结果中，第二匹

35

配结果中的“密码”也存在与第一匹配结果中，但是第一匹配结果中的“重新”并未存在与第二匹配结果中。

S604：当第二匹配结果中的每一字符均在第一匹配结果中存在相匹配的字符时，则返回第一匹配结果完全包含第二匹配结果的结果。

5 具体地，上述实施例中，当通过高层滤镜库匹配出的第一匹配结果与第二匹配结果完全相同，或者如上述例子中，两者唯一的区别点仅在于“重新”且该“重新”字符包含在第一匹配结果中，则认为第一匹配结果完全包含第二匹配结果。

S606：当第二匹配结果中存在至少一个字符在第一匹配结果中不存在相匹配的字符时，则返回第一匹配结果未完全包含第二匹配结果的结果。

10 具体地，当上述实施例中，通过高层滤镜库匹配出的第一匹配结果为“重新设置”，通过基础滤镜库匹配出的第二匹配结果为“账户异常，重新设置”时，即第一匹配结果与第二匹配结果仅“重新设置”字符相同，“账户异常”字符仅存在与第二匹配结果中，因此第一匹配结果未完全包含第二匹配结果。

15 上述实施例中，第一匹配结果是通过高层滤镜库进行匹配所得到的精确结果，第二匹配结果是通过基础滤镜库进行匹配所得到的普适结果，通过判断第一匹配结果和第二匹配结果的关系，来确定最终的目标匹配结果，提高了目标匹配结果的准确率。

应该理解的是，虽然图 2、图 3、图 5、图 6 的流程图中的各个步骤按照箭头的指示依次显示，但是这些步骤并不是必然按照箭头指示的顺序依次执行。除非本文中有明确的说明，这些步骤的执行并没有严格的顺序限制，这些步骤可以以其它的顺序执行。而且，
20 图 2、图 3、图 5、图 6 中的至少一部分步骤可以包括多个子步骤或者多个阶段，这些子步骤或者阶段并不必然是在同一时刻执行完成，而是可以在不同的时刻执行，这些子步骤或者阶段的执行顺序也不必然是依次进行，而是可以与其它步骤或者其它步骤的子步骤或者阶段的至少一部分轮流或者交替地执行。

参阅图 7，提供一种网页错误信息处理装置的结构示意图，该网页错误信息处理装置
25 包括：

爬取模块 100，用于爬取网页中的错误信息。

匹配模块 200，用于当标准库中不存在与错误信息相匹配的提示信息时，则将爬取的错误信息与高层滤镜库中的字符进行匹配以得到第一匹配结果；将爬取的错误信息与基础滤镜库中的字符进行匹配以得到第二匹配结果。

30 比较模块 300，用于将第一匹配结果与第二匹配结果进行比较；当第一匹配结果未完全包含第二匹配结果时，则将第一匹配结果和第二匹配结果进行拼接以获得目标匹配结果。

获取模块 400，用于根据目标匹配结果，获取爬取的错误信息的提示信息。

返回模块 500，用于将提示信息返回至网页。

35 在其中一个实施例中，获取模块可以包括：

接收单元, 用于接收针对目标匹配结果的调整指令。

处理单元, 用于当调整指令表示确认时, 则根据目标匹配结果, 获取爬取的错误信息的提示信息; 当调整指令表示调整时, 则根据调整指令对目标匹配结果进行调整, 并根据调整后的目标匹配结果, 获取爬取的错误信息的提示信息。

5 在其中一个实施例中, 该装置还可以包括:

补充模块, 用于将根据目标匹配结果获取的提示信息以及相应的错误信息关联存储至补充库。

匹配模块还用于当标准库中不存在与错误信息相匹配的提示信息时, 则将爬取的错误信息与补充库中已存储的错误信息进行匹配。

10 获取模块还用于当爬取的错误信息与补充库中已存储的错误信息匹配成功时, 则获取匹配成功的已存储的错误信息相对应的提示信息。

返回模块还用于将提示信息返回至网页。

比较模块还用于当爬取的错误信息与补充库中已存储的错误信息匹配失败时, 则继续将错误信息与高层滤镜库中的字符进行匹配以得到第一匹配结果。

15 在其中一个实施例中, 比较模块还可以用于当第一匹配结果完全包含第二匹配结果时, 则将第一匹配结果作为目标匹配结果。

在其中一个实施例中, 比较模块还可以用于当第一匹配结果中不存在字符时, 则将第二匹配结果作为目标匹配结果。

在其中一个实施例中, 比较模块可以包括:

20 匹配单元, 用于当第一匹配结果中存在字符时, 则通过第二匹配结果中的每一字符去匹配第一匹配结果中的每一字符。

输出单元, 用于当第二匹配结果中的每一字符均在第一匹配结果中存在相匹配的字符时, 则返回第一匹配结果完全包含第二匹配结果的结果; 当第二匹配结果中存在至少一个字符在第一匹配结果中不存在相匹配的字符时, 则返回第一匹配结果未完全包含第二匹配结果的结果。

25 上述对于网页错误信息处理装置的限定可以参见上文中对于坐席录屏方法的具体限定, 在此不再赘述。上述网页错误信息处理装置中的各个模块可全部或部分通过软件、硬件及其组合来实现。上述各模块可以硬件形式内嵌于或独立于计算机设备中的处理器中, 也可以以软件形式存储于计算机设备中的存储器中, 以便于处理器调用执行以上各个模块对应的操作。

30 在一个实施例中, 提供了一种计算机设备, 该计算机设备可以是服务器, 其内部结构图可以如图 8 所示。该计算机设备包括通过系统总线连接的处理器、存储器、网络接口和数据库。其中, 该计算机设备的处理器用于提供计算和控制能力。该计算机设备的存储器包括非易失性计算机可读指令存储介质、内存储器。该非易失性计算机可读指令存储介质存储有操作系统、计算机可读指令和数据库。该内存储器为非易失性计算机可读指令存储

35

介质中的操作系统和计算机可读指令的运行提供环境。该计算机设备的数据库用于存储xxx数据。该计算机设备的网络接口用于与外部的终端通过网络连接通信。该计算机可读指令被处理器执行时以实现一种网页错误信息处理方法。

本领域技术人员可以理解，图8中示出的结构，仅仅是与本申请方案相关的部分结构的框图，并不构成对本申请方案所应用于其上的计算机设备的限定，具体的计算机设备可以包括比图中所示更多或更少的部件，或者组合某些部件，或者具有不同的部件布置。

一种计算机设备，包括存储器和一个或多个处理器，存储器中储存有计算机可读指令，计算机可读指令被处理器执行时，使得一个或多个处理器执行以下步骤：爬取网页中的错误信息；当标准库中不存在与错误信息相匹配的提示信息时，则将爬取的错误信息与高层滤镜库中的字符进行匹配以得到第一匹配结果；将爬取的错误信息与基础滤镜库中的字符进行匹配以得到第二匹配结果；将第一匹配结果与第二匹配结果进行比较；当第一匹配结果未完全包含第二匹配结果时，则将第一匹配结果和第二匹配结果进行拼接以获得目标匹配结果；根据目标匹配结果，获取爬取的错误信息的提示信息；将提示信息返回至网页。

在其中一个实施例中，处理器执行计算机可读指令时还可以实现以下步骤：接收针对目标匹配结果的调整指令；当调整指令表示确认时，则根据目标匹配结果，获取爬取的错误信息的提示信息；当调整指令表示调整时，则根据调整指令对目标匹配结果进行调整，并根据调整后的目标匹配结果，获取爬取的错误信息的提示信息。

在其中一个实施例中，处理器执行计算机可读指令时还可以实现以下步骤：将根据目标匹配结果获取的提示信息以及相应的错误信息关联存储至补充库；当标准库中不存在与错误信息相匹配的提示信息时，则将爬取的错误信息与高层滤镜库中的字符进行匹配以得到第一匹配结果的步骤，包括：当标准库中不存在与错误信息相匹配的提示信息时，则将爬取的错误信息与补充库中已存储的错误信息进行匹配；当爬取的错误信息与补充库中已存储的错误信息匹配成功时，则获取匹配成功的已存储的错误信息相对应的提示信息；将提示信息返回至网页；当爬取的错误信息与补充库中已存储的错误信息匹配失败时，则继续将错误信息与高层滤镜库中的字符进行匹配以得到第一匹配结果的步骤。

在其中一个实施例中，处理器执行计算机可读指令时还可以实现以下步骤：当第一匹配结果完全包含第二匹配结果时，则将第一匹配结果作为目标匹配结果。

在其中一个实施例中，处理器执行计算机可读指令时还可以实现以下步骤：当第一匹配结果中不存在字符时，则将第二匹配结果作为目标匹配结果。

在其中一个实施例中，处理器执行计算机可读指令时还可以实现以下步骤：当第一匹配结果中存在字符时，则通过第二匹配结果中的每一字符去匹配第一匹配结果中的每一字符；当第二匹配结果中的每一字符均在第一匹配结果中存在相匹配的字符时，则返回第一匹配结果完全包含第二匹配结果的结果；当第二匹配结果中存在至少一个字符在第一匹配结果中不存在相匹配的字符时，则返回第一匹配结果未完全包含第二匹配结果的结果。

一个或多个存储有计算机可读指令的非易失性存储介质，计算机可读指令被一个或多个

个处理器执行时,使得一个或多个处理器执行以下步骤:爬取网页中的错误信息;当标准库中不存在与错误信息相匹配的提示信息时,则将爬取的错误信息与高层滤镜库中的字符进行匹配以得到第一匹配结果;将爬取的错误信息与基础滤镜库中的字符进行匹配以得到第二匹配结果;将第一匹配结果与第二匹配结果进行比较;当第一匹配结果未完全包含第二匹配结果时,则将第一匹配结果和第二匹配结果进行拼接以获得目标匹配结果;根据目标匹配结果,获取爬取的错误信息的提示信息;将提示信息返回至网页。

在其中一个实施例中,该计算机可读指令被处理器执行时还可以实现以下步骤:接收针对目标匹配结果的调整指令;当调整指令表示确认时,则根据目标匹配结果,获取爬取的错误信息的提示信息;当调整指令表示调整时,则根据调整指令对目标匹配结果进行调整,并根据调整后的目标匹配结果,获取爬取的错误信息的提示信息。

在其中一个实施例中,该计算机可读指令被处理器执行时还可以实现以下步骤:将根据目标匹配结果获取的提示信息以及相应的错误信息关联存储至补充库;当标准库中不存在与错误信息相匹配的提示信息时,则将爬取的错误信息与高层滤镜库中的字符进行匹配以得到第一匹配结果的步骤,包括:当标准库中不存在与错误信息相匹配的提示信息时,则将爬取的错误信息与补充库中已存储的错误信息进行匹配;当爬取的错误信息与补充库中已存储的错误信息匹配成功时,则获取匹配成功的已存储的错误信息相对应的提示信息;将提示信息返回至网页;当爬取的错误信息与补充库中已存储的错误信息匹配失败时,则继续将错误信息与高层滤镜库中的字符进行匹配以得到第一匹配结果的步骤。

在其中一个实施例中,该计算机可读指令被处理器执行时还可以实现以下步骤:当第一匹配结果完全包含第二匹配结果时,则将第一匹配结果作为目标匹配结果。

在其中一个实施例中,该计算机可读指令被处理器执行时还可以实现以下步骤:当第一匹配结果中不存在字符时,则将第二匹配结果作为目标匹配结果。

在其中一个实施例中,该计算机可读指令被处理器执行时还可以实现以下步骤:当第一匹配结果中存在字符时,则通过第二匹配结果中的每一字符去匹配第一匹配结果中的每一字符;当第二匹配结果中的每一字符均在第一匹配结果中存在相匹配的字符时,则返回第一匹配结果完全包含第二匹配结果的结果;当第二匹配结果中存在至少一个字符在第一匹配结果中不存在相匹配的字符时,则返回第一匹配结果未完全包含第二匹配结果的结果。

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,是可以通过计算机可读指令来指令相关的硬件来完成,所述的计算机可读指令可存储于一非易失性计算机可读存储介质中,该计算机可读指令在执行时,可包括如上述各方法的实施例的流程。其中,本申请所提供的各实施例中所使用的对存储器、存储、数据库或其它介质的任何引用,均可包括非易失性和/或易失性存储器。非易失性存储器可包括只读存储器(ROM)、可编程ROM(PROM)、电可编程ROM(EPROM)、电可擦除可编程ROM(EEPROM)或闪存。易失性存储器可包括随机存取存储器(RAM)或者外部高速缓冲

存储器。作为说明而非局限，RAM 以多种形式可得，诸如静态 RAM (SRAM)、动态 RAM (DRAM)、同步 DRAM (SDRAM)、双数据率 SDRAM (DDRSDRAM)、增强型 SDRAM (ESDRAM)、同步链路 (Synchlink) DRAM (SLDRAM)、存储器总线 (Rambus) 直接 RAM (RDRAM)、直接存储器总线动态 RAM (DRDRAM)、以及存储器总线动态 RAM (RDRAM) 等。

以上所述实施例的各技术特征可以进行任意的组合，为使描述简洁，未对上述实施例中的各个技术特征所有可能的组合都进行描述，然而，只要这些技术特征的组合不存在矛盾，都应当认为是本说明书记载的范围。

以上所述实施例仅表达了本申请的几种实施方式，其描述较为具体和详细，但并不能因此而理解为对发明专利范围的限制。应当指出的是，对于本领域的普通技术人员来说，在不脱离本申请构思的前提下，还可以做出若干变形和改进，这些都属于本申请的保护范围。因此，本申请专利的保护范围应以所附权利要求为准。

权利要求书

1、一种网页错误信息处理方法，包括：

爬取网页中的错误信息；

当标准库中不存在与所述错误信息相匹配的提示信息时，则将爬取的所述错误信息与高层滤镜库中的字符进行匹配以得到第一匹配结果；

5 将爬取的所述错误信息与基础滤镜库中的字符进行匹配以得到第二匹配结果；

将所述第一匹配结果与所述第二匹配结果进行比较；

当所述第一匹配结果未完全包含所述第二匹配结果时，则将第一匹配结果和所述第二匹配结果进行拼接以获得目标匹配结果；

根据所述目标匹配结果，获取爬取的所述错误信息的提示信息；及

10 将所述提示信息返回至所述网页。

2、根据权利要求 1 所述的方法，其特征在于，所述根据所述目标匹配结果，获取爬取的所述错误信息的提示信息，包括：

接收针对所述目标匹配结果的调整指令；

15 当所述调整指令表示确认时，则根据所述目标匹配结果，获取爬取的所述错误信息的提示信息；及

当所述调整指令表示调整时，则根据所述调整指令对所述目标匹配结果进行调整，并根据调整后的所述目标匹配结果，获取爬取的所述错误信息的提示信息。

3、根据权利要求 2 所述的方法，其特征在于，所述方法还包括：

20 将根据所述目标匹配结果获取的提示信息以及相应的所述错误信息关联存储至补充库；

所述当标准库中不存在与所述错误信息相匹配的提示信息时，则将爬取的所述错误信息与高层滤镜库中的字符进行匹配以得到第一匹配结果，包括：

当标准库中不存在与所述错误信息相匹配的提示信息时，则将爬取的所述错误信息与所述补充库中已存储的错误信息进行匹配；

25 当爬取的所述错误信息与所述补充库中已存储的错误信息匹配成功时，则获取匹配成功的已存储的错误信息相对应的提示信息；

将所述提示信息返回至所述网页；及

当爬取的所述错误信息与所述补充库中已存储的错误信息匹配失败时，则继续将所述错误信息与高层滤镜库中的字符进行匹配以得到第一匹配结果的步骤。

30 4、根据权利要求 1 至 3 任一项所述的方法，其特征在于，所述方法还包括：

当所述第一匹配结果完全包含所述第二匹配结果时，则将所述第一匹配结果作为所述目标匹配结果。

5、根据权利要求 4 所述的方法，其特征在于，所述方法还包括：

当所述第一匹配结果中不存在字符时，则将所述第二匹配结果作为所述目标匹配结

果。

6、根据权利要求 4 所述的方法，其特征在于，所述将所述第一匹配结果与所述第二匹配结果进行比较，包括：

5 当所述第一匹配结果中存在字符时，则通过所述第二匹配结果中的每一字符去匹配所述第一匹配结果中的每一字符；

当所述第二匹配结果中的每一字符均在所述第一匹配结果中存在相匹配的字符时，则返回所述第一匹配结果完全包含所述第二匹配结果的结果；及

当所述第二匹配结果中存在至少一个字符在所述第一匹配结果中不存在相匹配的字符时，则返回所述第一匹配结果未完全包含所述第二匹配结果的结果。

10

7、一种网页错误信息处理装置，包括：

爬取模块，用于爬取网页中的错误信息；

15 匹配模块，用于当标准库中不存在与所述错误信息相匹配的提示信息时，则将爬取的所述错误信息与高层滤镜库中的字符进行匹配以得到第一匹配结果；将爬取的所述错误信息与基础滤镜库中的字符进行匹配以得到第二匹配结果；

比较模块，用于将所述第一匹配结果与所述第二匹配结果进行比较；当所述第一匹配结果未完全包含所述第二匹配结果时，则将第一匹配结果和所述第二匹配结果进行拼接以获得目标匹配结果；

20 获取模块，用于根据所述目标匹配结果，获取爬取的所述错误信息的提示信息；及
返回模块，用于将所述提示信息返回至所述网页。

8、根据权利要求 7 所述的装置，其特征在于，所述获取模块包括：

接收单元，用于接收针对所述目标匹配结果的调整指令；及

25 处理单元，用于当所述调整指令表示确认时，则根据所述目标匹配结果，获取爬取的所述错误信息的提示信息；当所述调整指令表示调整时，则根据所述调整指令对所述目标匹配结果进行调整，并根据调整后的所述目标匹配结果，获取爬取的所述错误信息的提示信息。

9、一种计算机设备，包括存储器及一个或多个处理器，所述存储器中储存有计算机可读指令，所述计算机可读指令被所述一个或多个处理器执行时，使得所述一个或多个处理器执行以下步骤：

30 爬取网页中的错误信息；

当标准库中不存在与所述错误信息相匹配的提示信息时，则将爬取的所述错误信息与高层滤镜库中的字符进行匹配以得到第一匹配结果；

将爬取的所述错误信息与基础滤镜库中的字符进行匹配以得到第二匹配结果；

将所述第一匹配结果与所述第二匹配结果进行比较；

35 当所述第一匹配结果未完全包含所述第二匹配结果时，则将第一匹配结果和所述第二

匹配结果进行拼接以获得目标匹配结果;

根据所述目标匹配结果, 获取爬取的所述错误信息的提示信息; 及
将所述提示信息返回至所述网页。

10、根据权利要求 9 所述的计算机设备, 其特征在于, 所述处理器执行所述计算机可
5 读指令时实现的所述根据所述目标匹配结果, 获取爬取的所述错误信息的提示信息, 包括:
接收针对所述目标匹配结果的调整指令;

当所述调整指令表示确认时, 则根据所述目标匹配结果, 获取爬取的所述错误信息的
提示信息; 及

10 当所述调整指令表示调整时, 则根据所述调整指令对所述目标匹配结果进行调整, 并
根据调整后的所述目标匹配结果, 获取爬取的所述错误信息的提示信息。

11、根据权利要求 10 所述的计算机设备, 其特征在于, 所述处理器执行所述计算机
可读指令时还执行以下步骤:

将根据所述目标匹配结果获取的提示信息以及相应的所述错误信息关联存储至补充
库;

15 所述当标准库中不存在与所述错误信息相匹配的提示信息时, 则将爬取的所述错误信
息与高层滤镜库中的字符进行匹配以得到第一匹配结果, 包括:

当标准库中不存在与所述错误信息相匹配的提示信息时, 则将爬取的所述错误信息与
所述补充库中已存储的错误信息进行匹配;

20 当爬取的所述错误信息与所述补充库中已存储的错误信息匹配成功时, 则获取匹配成
功的已存储的错误信息相对应的提示信息;

将所述提示信息返回至所述网页; 及

当爬取的所述错误信息与所述补充库中已存储的错误信息匹配失败时, 则继续将所述
错误信息与高层滤镜库中的字符进行匹配以得到第一匹配结果的步骤。

12、根据权利要求 9 至 11 任一项所述的计算机设备, 其特征在于, 所述处理器执行
25 所述计算机可读指令时还执行以下步骤:

当所述第一匹配结果完全包含所述第二匹配结果时, 则将所述第一匹配结果作为所述
目标匹配结果。

13、根据权利要求 12 所述的计算机设备, 其特征在于, 所述处理器执行所述计算机
可读指令时还执行以下步骤:

30 当所述第一匹配结果中不存在字符时, 则将所述第二匹配结果作为所述目标匹配结
果。

14、根据权利要求 12 所述的计算机设备, 其特征在于, 所述处理器执行所述计算机
可读指令时实现的所述将所述第一匹配结果与所述第二匹配结果进行比较, 包括:

35 当所述第一匹配结果中存在字符时, 则通过所述第二匹配结果中的每一字符去匹配所
述第一匹配结果中的每一字符;

当所述第二匹配结果中的每一字符均在所述第一匹配结果中存在相匹配的字符时，则返回所述第一匹配结果完全包含所述第二匹配结果的结果；及

当所述第二匹配结果中存在至少一个字符在所述第一匹配结果中不存在相匹配的字符时，则返回所述第一匹配结果未完全包含所述第二匹配结果的结果。

5 15、一个或多个存储有计算机可读指令的非易失性计算机可读存储介质，所述计算机可读指令被一个或多个处理器执行时，使得所述一个或多个处理器执行以下步骤：

爬取网页中的错误信息；

当标准库中不存在与所述错误信息相匹配的提示信息时，则将爬取的所述错误信息与高层滤镜库中的字符进行匹配以得到第一匹配结果；

10 将爬取的所述错误信息与基础滤镜库中的字符进行匹配以得到第二匹配结果；

将所述第一匹配结果与所述第二匹配结果进行比较；

当所述第一匹配结果未完全包含所述第二匹配结果时，则将第一匹配结果和所述第二匹配结果进行拼接以获得目标匹配结果；

根据所述目标匹配结果，获取爬取的所述错误信息的提示信息；及

15 将所述提示信息返回至所述网页。

16、根据权利要求 15 所述的存储介质，其特征在于，所述计算机可读指令被所述处理器执行时实现的所述根据所述目标匹配结果，获取爬取的所述错误信息的提示信息，包括：

接收针对所述目标匹配结果的调整指令；

20 当所述调整指令表示确认时，则根据所述目标匹配结果，获取爬取的所述错误信息的提示信息；及

当所述调整指令表示调整时，则根据所述调整指令对所述目标匹配结果进行调整，并根据调整后的所述目标匹配结果，获取爬取的所述错误信息的提示信息。

25 17、根据权利要求 16 所述的存储介质，其特征在于，所述计算机可读指令被所述处理器执行时还执行以下步骤：

将根据所述目标匹配结果获取的提示信息以及相应的所述错误信息关联存储至补充库；

所述当标准库中不存在与所述错误信息相匹配的提示信息时，则将爬取的所述错误信息与高层滤镜库中的字符进行匹配以得到第一匹配结果，包括：

30 当标准库中不存在与所述错误信息相匹配的提示信息时，则将爬取的所述错误信息与所述补充库中已存储的错误信息进行匹配；

当爬取的所述错误信息与所述补充库中已存储的错误信息匹配成功时，则获取匹配成功的已存储的错误信息相对应的提示信息；

将所述提示信息返回至所述网页；及

35 当爬取的所述错误信息与所述补充库中已存储的错误信息匹配失败时，则继续将所述

错误信息与高层滤镜库中的字符进行匹配以得到第一匹配结果的步骤。

18、根据权利要求 15 至 17 任一项所述的存储介质，其特征在于，所述计算机可读指令被所述处理器执行时还执行以下步骤：

5 当所述第一匹配结果完全包含所述第二匹配结果时，则将所述第一匹配结果作为所述目标匹配结果。

19、根据权利要求 18 所述的存储介质，其特征在于，所述计算机可读指令被所述处理器执行时还执行以下步骤：

当所述第一匹配结果中不存在字符时，则将所述第二匹配结果作为所述目标匹配结果。

10 20、根据权利要求 18 所述的存储介质，其特征在于，所述计算机可读指令被所述处理器执行时实现的所述将所述第一匹配结果与所述第二匹配结果进行比较，包括：

当所述第一匹配结果中存在字符时，则通过所述第二匹配结果中的每一字符去匹配所述第一匹配结果中的每一字符；

15 当所述第二匹配结果中的每一字符均在所述第一匹配结果中存在相匹配的字符时，则返回所述第一匹配结果完全包含所述第二匹配结果的结果；及

当所述第二匹配结果中存在至少一个字符在所述第一匹配结果中不存在相匹配的字符时，则返回所述第一匹配结果未完全包含所述第二匹配结果的结果。

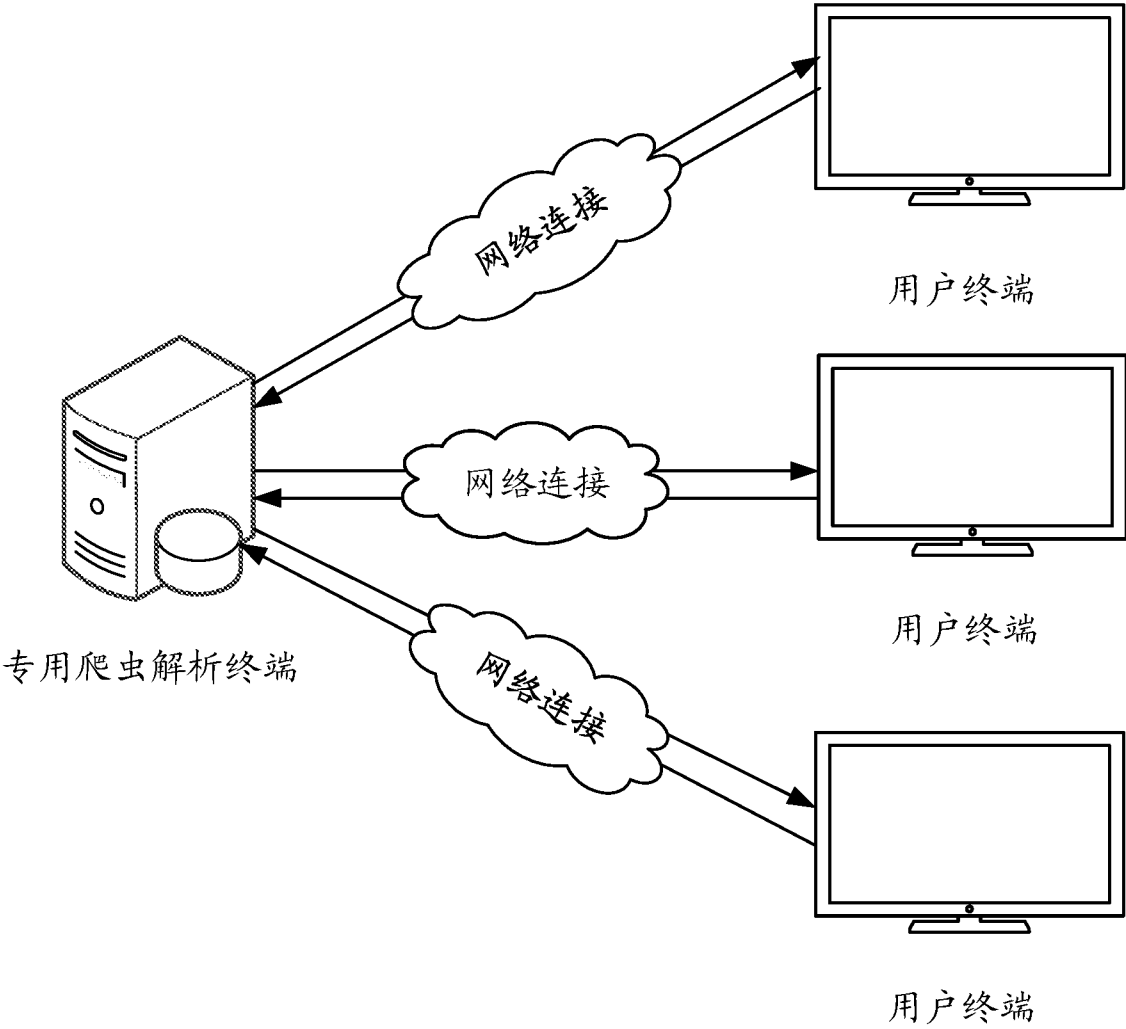


图 1

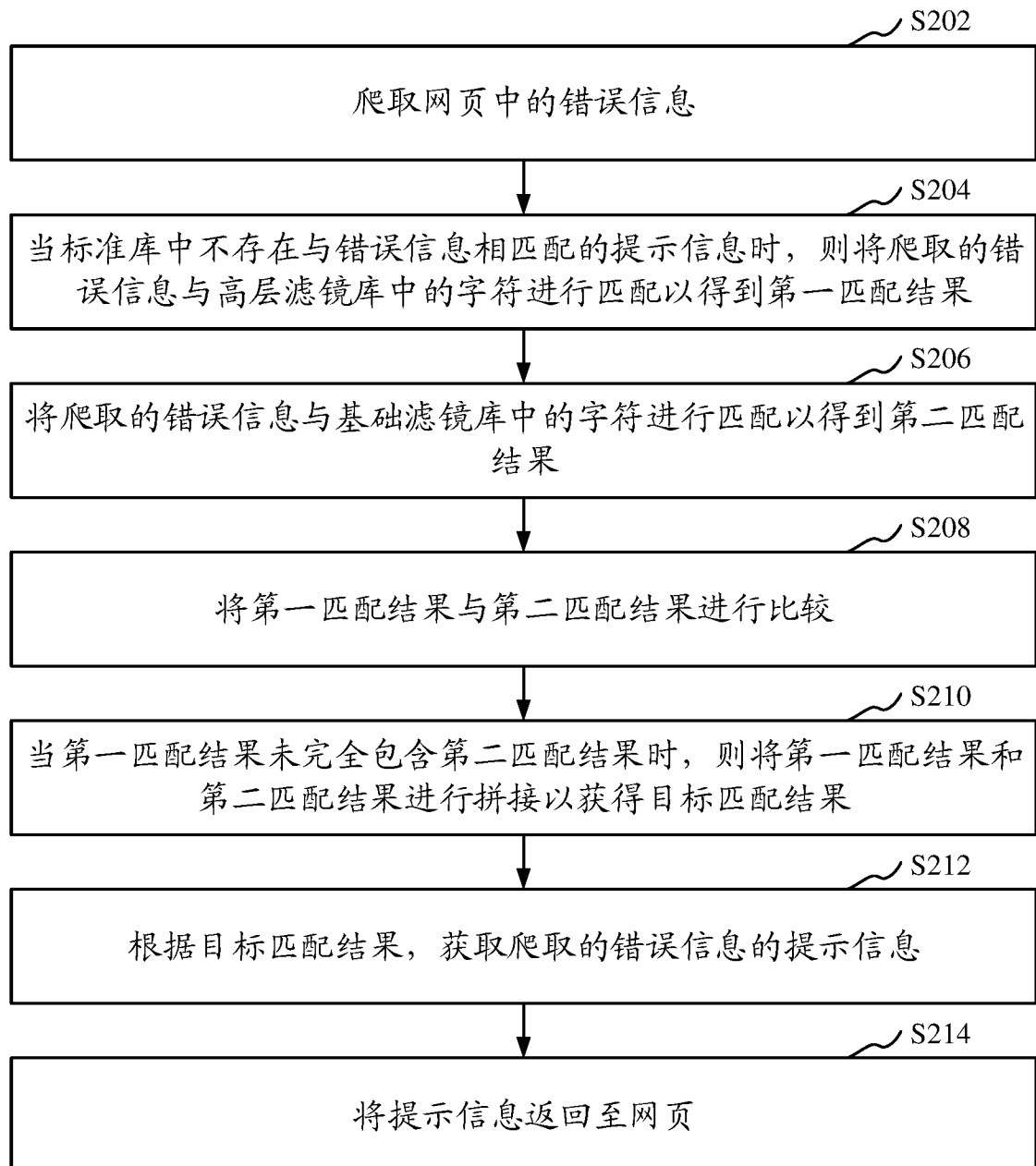


图 2

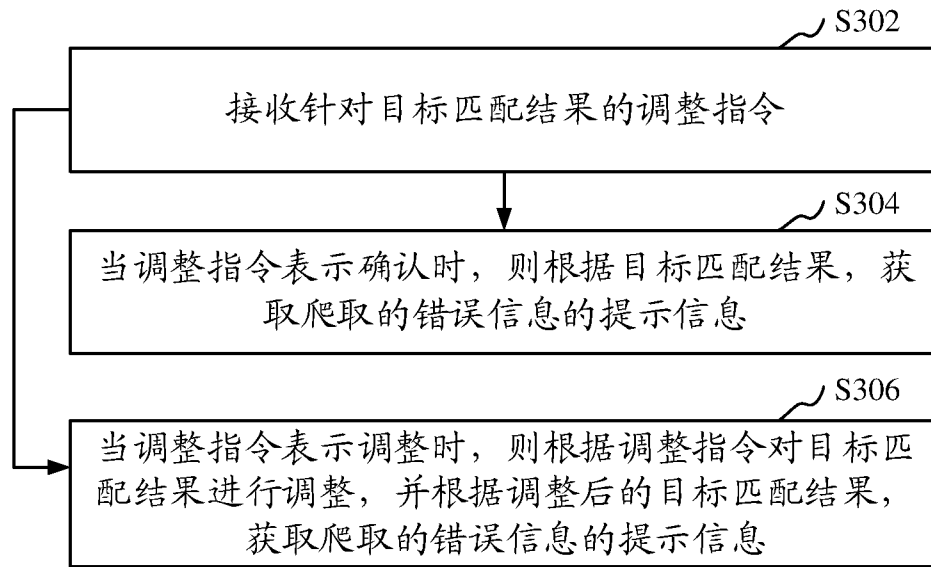


图 3

原始错误信息	匹配码	目标匹配结果	确认	调整	撤回
账户异常，请去柜台重新设置	4217	账户异常重新设置	☒		

图 4

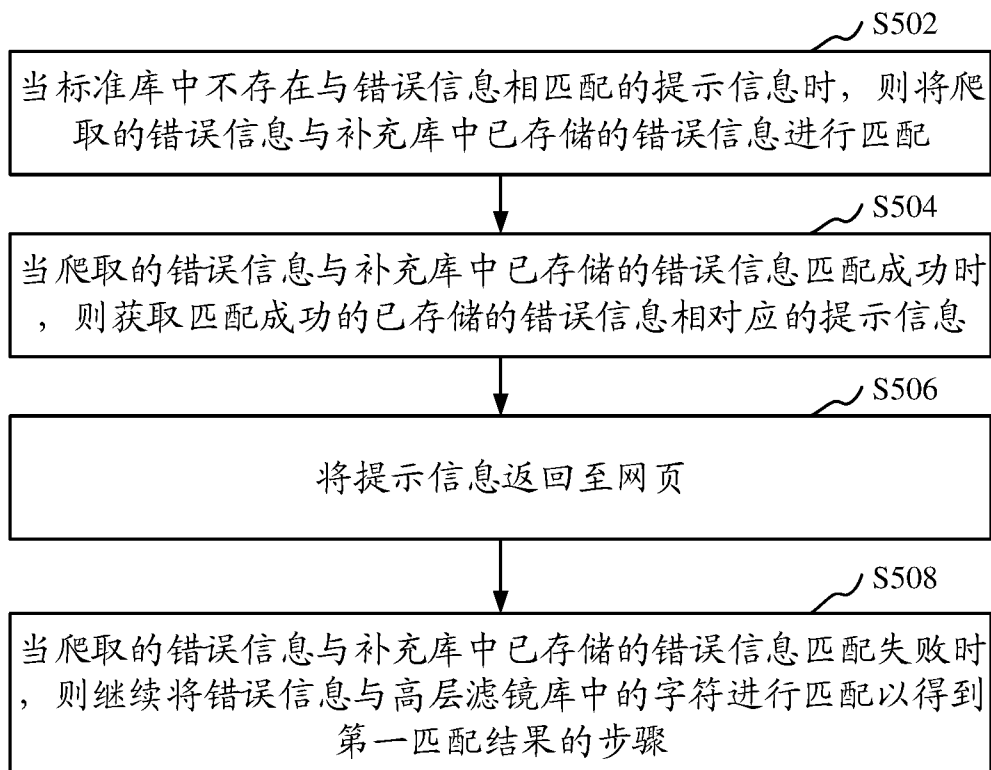


图 5

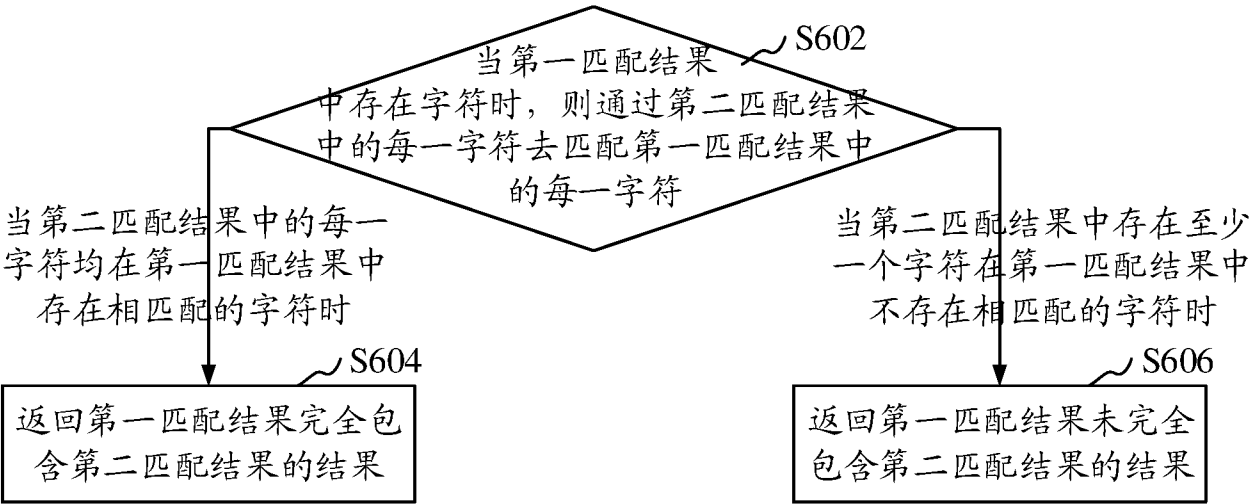


图 6

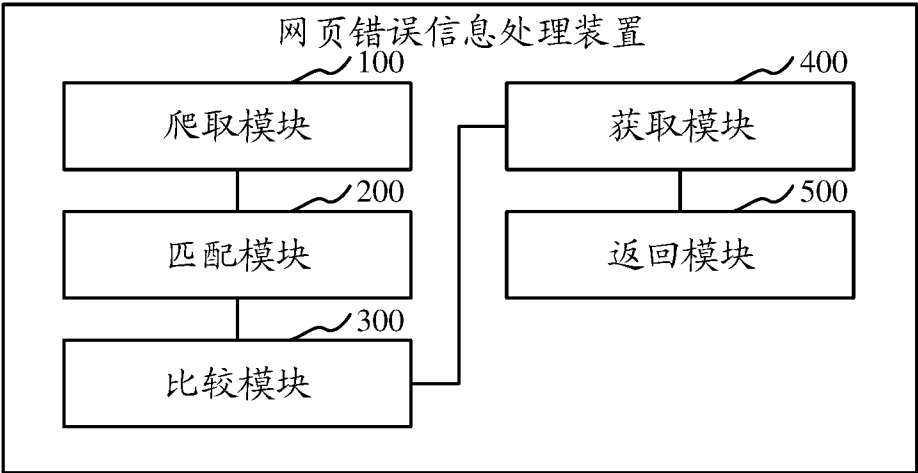


图 7

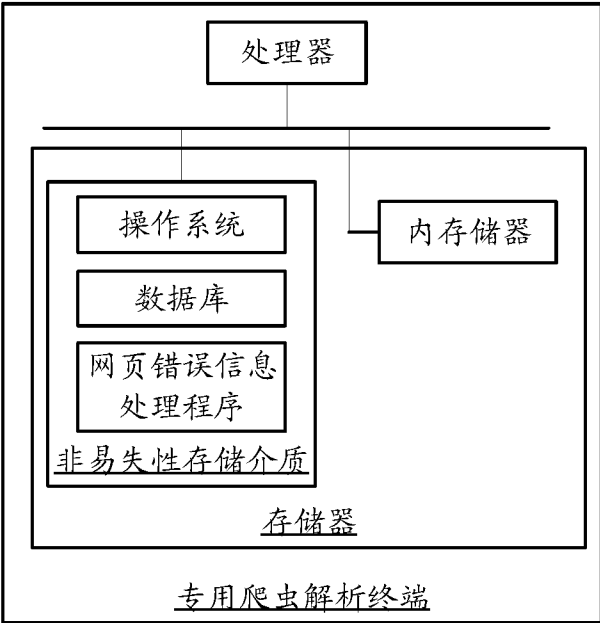


图 8

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2018/080139

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)

CNKI, CNPAT, WPI, EPODOC, IEEE: 网页, 错误, 报错, 出错, 信息, 匹配, 提示, 拼接, 合并, 规则库, 知识库, 规则, 结果, 比较, 转换, 映射; webpage, error, message, information, prompt, map+, match+, convert+, rule, knowledge, library, compar+, splic+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 107704484 A (SHANGHAI YIZHANGTONG FINANCIAL TECHNOLOGY CO., LTD.) 16 February 2018 (16.02.2018), claims 1-10	1-20
A	CN 104750747 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 01 July 2015 (01.07.2015), description, paragraphs [0048]-[0054], [0073]-[0082], [0147]-[0150] and [0175], and figures 1, 2 and 4	1-20
A	CN 102306177 A (TSINGHUA UNIVERSITY) 04 January 2012 (04.01.2012), entire document	1-20
A	CN 106648944 A (NEUSOFT GROUP CORP.) 10 May 2017 (10.05.2017), entire document	1-20
A	CN 105550059 A (CHINA CONSTRUCTION BANK CO., LTD.) 04 May 2016 (04.05.2016), entire document	1-20

☒ 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 13 June 2018	Date of mailing of the international search report 26 June 2018
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 CAO, Shumei Telephone No. (86-10) 53961315

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2018/080139

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2008147652 A1 (BELLSOUTH INTELLECTUAL PROPERTY CORPORATION) 19 June 2008 (19.06.2008), entire document	1-20

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2018/080139

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 107704484 A	16 February 2018	None	
CN 104750747 A	01 July 2015	WO 2015101143 A1	09 July 2015
CN 102306177 A	04 January 2012	None	
CN 106648944 A	10 May 2017	None	
CN 105550059 A	04 May 2016	None	
US 2008147652 A1	19 June 2008	None	

国际检索报告

国际申请号

PCT/CN2018/080139

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))

CNKI, CNPAT, WPI, EPODOC, IEEE: 网页, 错误, 报错, 出错, 信息, 匹配, 提示, 拼接, 合并, 规则库, 知识库, 规则, 结果, 比较, 转换, 映射; webpage, error, message, information, prompt, map+, match+, convert+, rule, knowledge, library, compar+, splic+

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
PX	CN 107704484 A (上海壹账通金融科技有限公司) 2018年 2月 16日 (2018 - 02 - 16) 权利要求1-10	1-20
A	CN 104750747 A (腾讯科技深圳有限公司) 2015年 7月 1日 (2015 - 07 - 01) 说明书第0048-0054、0073-0082、0147-0150、0175段, 图1、2、4	1-20
A	CN 102306177 A (清华大学) 2012年 1月 4日 (2012 - 01 - 04) 全文	1-20
A	CN 106648944 A (东软集团股份有限公司) 2017年 5月 10日 (2017 - 05 - 10) 全文	1-20
A	CN 105550059 A (中国建设银行股份有限公司) 2016年 5月 4日 (2016 - 05 - 04) 全文	1-20
A	US 2008147652 A1 (BELLSOUTH INTELLECTUAL PROPERTY CORPORATION) 2008年 6月 19日 (2008 - 06 - 19) 全文	1-20

☐ 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

2018年 6月 13日

国际检索报告邮寄日期

2018年 6月 26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

受权官员

曹姝妹

电话号码 86-(10)-53961315

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2018/080139

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	107704484	A	2018年 2月 16日	无			
CN	104750747	A	2015年 7月 1日	WO	2015101143	A1	2015年 7月 9日
CN	102306177	A	2012年 1月 4日	无			
CN	106648944	A	2017年 5月 10日	无			
CN	105550059	A	2016年 5月 4日	无			
US	2008147652	A1	2008年 6月 19日	无			