(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
2020 年 1 月 2 日 (02.01.2020)

WIPO | PCT

(10) 国际公布号
WO 2020/000747 A1

(51) 国际专利分类号:
G06F 21/54 (2013.01)

(21) 国际申请号: PCT/CN2018/108672

(22) 国际申请日: 2018 年 9 月 29 日 (29.09.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201810685659.3 2018年6月27日 (27.06.2018) CN

(71) 申请人: 平安科技（深圳）有限公司 (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN]; 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。

(72) 发明人: 邵壮 (SHAO, Zhuang); 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。

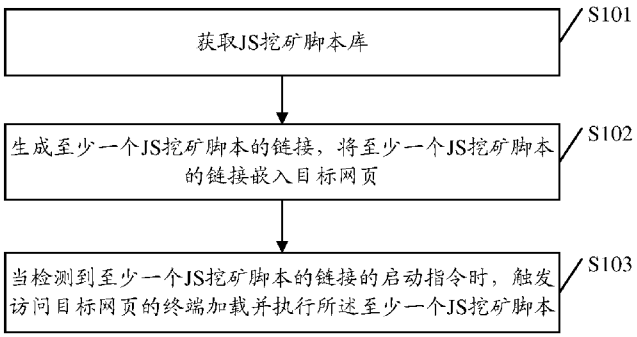
(74) 代理人: 广州三环专利商标代理有限公司 (SCIHEAD IP LAW FIRM); 中国广东省广州市越秀区先烈中路80号汇华商贸大厦1508室, Guangdong 510070 (CN)。

(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, 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,

(54) **Title:** ANTI-CRAWLER METHOD AND TERMINAL AND COMPUTER READABLE STORAGE MEDIUM

(54) **发明名称:** 一种反爬虫方法、终端及计算机可读存储介质



S101 ACQUIRE A JS MINING SCRIPT LIBRARY
S102 GENERATE A LINK FOR AT LEAST ONE JS MINING SCRIPT, AND EMBED THE LINK OF THE AT LEAST ONE JS MINING SCRIPT IN A TARGET WEBPAGE
S103 WHEN A START-UP INSTRUCTION FOR THE LINK OF THE AT LEAST ONE JS MINING SCRIPT IS DETECTED, TRIGGER A TERMINAL WHICH ACCESSES SAID TARGET WEBPAGE TO LOAD AND RUN THE AT LEAST ONE JS MINING SCRIPT

图 1

(57) **Abstract:** An anti-crawler method and terminal and a computer readable storage medium, the method comprising: acquiring a JS mining script library (101), the JS mining script library comprising at least one JS mining script; generating a link for at least one JS mining script, and embedding the link of the at least one JS mining script in a target webpage (S102); when a start-up instruction for the link of the at least one JS mining script is detected, triggering a terminal which accesses said target webpage to load and run the at least one JS mining script (S103). The described method may greatly consume the CPU resources of a crawler terminal so that the crawler terminal may not be normally used and the purpose of preventing crawling may be achieved; and the method may also prevent crawler terminals from bypassing JS and rendering crawling data, thus increasing the effectiveness of preventing crawling.

WO 2020/000747 A1

[见续页]

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))。

(57) 摘要: 一种反爬虫方法、终端及计算机可读存储介质，其中方法包括：获取JS挖矿脚本库(S101)，所述JS挖矿脚本库包括至少一个JS挖矿脚本；生成所述至少一个JS挖矿脚本的链接，将所述至少一个JS挖矿脚本的链接嵌入目标网页(S102)；当检测到所述至少一个JS挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端加载并执行所述至少一个JS挖矿脚本(S103)。上述方法可以大量消耗爬虫终端的CPU资源，使爬虫终端无法正常使用，以达到反爬的目的，以及可以防止爬虫终端绕过JS渲染爬取数据，提高了反爬的有效性。

一种反爬虫方法、终端及计算机可读存储介质

本申请要求于 2018 年 06 月 27 日提交中国专利局、申请号为 201810685659.3、申请名称为“一种反爬虫方法、终端及计算机可读介质”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

本申请涉及通信技术领域，尤其涉及一种反爬虫方法、终端及计算机可读存储介质。

背景技术

目前互联网上的爬虫主要分为静态爬虫和动态爬虫，由于静态爬虫不能解析 JavaScript （简称 JS）代码，因此对于静态爬虫的应对方法通常是采用通过在网页中添加 JS 代码，以对网页进行 JS 渲染处理的方式来达到反爬的目的。然而这种对网页进行 JS 渲染处理的方式对于能够解析 JS 的动态爬虫不能达到反爬虫的目的。因此，如何更有效地阻止爬虫的影响，提高互联网的安全成为研究的热点。

发明内容

本申请实施例提供一种反爬虫方法、终端及计算机可读存储介质，可提高反爬的有效性，提高互联网安全性。

第一方面，本申请实施例提供了一种反爬虫方法，该方法包括:

获取 JS 挖矿脚本库，所述 JS 挖矿脚本库包括至少一个 JS 挖矿脚本;

生成所述至少一个 JS 挖矿脚本的链接，将所述至少一个 JS 挖矿脚本的链接嵌入目标网页;

当检测到所述至少一个 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端加载并执行所述至少一个 JS 挖矿脚本。

第二方面，本申请实施例提供了一种终端，该终端包括用于执行上述第一方面的方法的单元。

第三方面，本申请实施例提供了另一种终端，包括处理器、输入设备、输出设备和存储器，所述处理器、输入设备、输出设备和存储器相互连接，其中，所述存储器用于存储支持终端执行上述方法的计算机程序，所述计算机程序包括程序指令，所述处理器被配置用于调用所述程序指令，执行上述第一方面的方法。

第四方面，本申请实施例提供了一种计算机可读存储介质，所述计算机存储介质存储有计算机程序，所述计算机程序包括程序指令，所述程序指令当被处理器执行时使所述处理器执行上述第一方面的方法。

本申请实施例，当检测到嵌入目标网页的 JS 挖矿脚本的链接的启动指令时，触发访问该目标网页的终端加载并执行该 JS 挖矿脚本，防止了爬虫终端绕过 JS 渲染爬取数据，提高了反爬的有效性。

附图说明

图 1 是本申请实施例提供的一种反爬虫方法的示意流程图；

图 2 是本申请实施例提供的另一种反爬虫方法的示意流程图；

图 3 是本申请实施例提供的又一种反爬虫方法的示意流程图；

图 4 是本申请实施例提供的一种终端的示意框图；

图 5 是本申请实施例提供的另一种终端示意框图。

具体实施方式

下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

本申请实施例提供的反爬虫方法可以由终端执行，所述终端可以是手机、电脑、平板、智能手表等智能终端上。下面对应用于终端的反爬虫方法进行举例说明。

本方案是针对能够解析 JS 的动态爬虫所做的反爬虫技术方案，现有技术中关于网页的常用反爬虫策略通常为在网页中添加 JS 代码来进行 JS 渲染处理，以实现对静态爬虫进行反爬，然而这种方式对于能够解析 JS 的动态爬虫而言，不能达到反爬虫的目的，因此本方案针对能够解析 JS 的动态爬虫对反爬虫策略进行了改进。

在介绍本方案之前，首先介绍一下能够解析 JS 的动态爬虫。能够解析 JS 的动态爬虫的爬虫过程有两种情况：一种是通过直接对 html 源码进行解析；另外还有一种特殊情况是：如果网页的数据是采用异步加载的，比如 ajax 加载的数据，则能够解析 JS 的动态爬虫无法解析并查看源代码，针对这种特殊情况，能够解析 JS 的动态爬虫常采用一些第三方的工具比如：支持 WebAssembly 语言的浏览器中的自动测试工具 Selenium 或者配合无界面浏览器 PhantomJs，模拟浏览器的行为去加载数据。需要说明的是，该 WebAssembly 是一种运行在现代网络浏览器中的一种网页汇编语言，文件后缀为“.wasm”。

因此，本方案针对能够解析 JS 的动态爬虫在爬虫过程中所使用的上述两种爬虫方式，预先编写了至少一个 JS 挖矿脚本，并建立 JS 挖矿脚本库，根据该 JS 挖矿脚本库中预先编写的至少一个 JS 挖矿脚本，生成至少一个 JS 挖矿脚本的链接。本申请实施例将生成的该至少一个 JS 挖矿脚本的链接嵌入目标

网页中数据信息被爬取的概率较高的位置区域，如果检测到访问该目标网页的终端打开嵌入该目标网页中的所述至少一个 JS 挖矿脚本的链接，则确定访问该目标网页的终端为爬虫终端，并触发访问该目标网页的终端自动加载和执行所述链接中的至少一个 JS 挖矿脚本，以达到反爬的目的。下面结合附图对本申请实施例进行详细说明。

参见图 1，图 1 是本申请实施例提供的一种反爬虫方法的示意流程图，如图 1 所示，该方法可以由终端执行，所述终端的具体解释如前所述，此处不再赘述。具体地，本申请实施例的所述方法包括如下步骤。

S101：获取 JS 挖矿脚本库。

本申请实施例中，终端可以建立 JS 挖矿脚本库，其中，该终端所建立的 JS 挖矿脚本库中包括一个或多个 JS 挖矿脚本。

在一个实施例中，该终端建立的 JS 挖矿脚本库中包括第一 JS 挖矿脚本和第二 JS 挖矿脚本。其中，所述第一 JS 挖矿脚本是 WebAssembly 语言的脚本，所述第二 JS 挖矿脚本是 JS 语言的脚本。

其中，所述第一 JS 挖矿脚本是根据 WebAssembly 语言编写的脚本，在一个实施例中，由于通常情况下能够解析 JS 的动态爬虫常使用 Selenium 配合 Webdriver 调用浏览器来模拟浏览器操作，其中，使用 Selenium 配合 Webdriver 调用的浏览器是支持该 WebAssembly 的浏览器；或者使用 Selenium 配合 PhantomJS 这种无界面的浏览器来模拟浏览器操作。因此本方案针对能够解析 JS 的动态爬虫常使用的支持该 WebAssembly 的浏览器或者 PhantomJS，利用 WebAssembly 语言编写了所述第一 JS 挖矿脚本。

在其他实施例中，所述第二 JS 挖矿脚本是 JS 语言的脚本，由于能够解析 JS 的动态爬虫还可以使用不支持 WebAssembly 的普通浏览器直接解析 html 源码进行解析，因此利用 JS 语言编写了第二 JS 挖矿脚本。

S102：生成至少一个 JS 挖矿脚本的链接，将至少一个 JS 挖矿脚本的链接嵌入目标网页。

本申请实施例中，终端可以生成至少一个 JS 挖矿脚本的链接，并将所述至少一个 JS 挖矿脚本的链接嵌入目标网页。

在一个实施例中，该终端可以根据由第一 JS 挖矿脚本和第二 JS 挖矿脚本建立的 JS 挖矿脚本库生成至少一个 JS 挖矿脚本的链接，并将所述至少一个 JS 挖矿脚本的连接嵌入目标网页的第一位置区域。

其中，所述第一位置区域是根据历史记录中记录的所述目标网页中各位置区域的数据信息被爬取的概率确定的，如果从所述目标网页的各位置区域中确定出某位置区域的数据信息被爬取的概率最大，则确定所述数据信息被爬取概率最大的位置区域为第一位置区域。例如，假设所述目标网页中包括评论位置区域、提问位置区域、图像显示位置区域、信息显示位置区域，如果获取到历史记录中记录了评论位

置区域的数据信息被爬取的概率为 90%，提问位置区域的数据信息被爬取的概率为 60%，图像显示位置区域的数据信息被爬取的概率为 20%，信息显示位置区域的数据信息被爬取的概率为 50%，将各位置区域的数据信息被爬取的概率进行对比，确定所述评论位置区域的数据信息被爬取的概率 90%最大，因此确定所述目标网页中的评论位置区域为第一位置区域。

在一个实施例中，所述第一位置区域可以根据历史记录中记录的所述目标网页中各位置区域的数据信息被爬取的关键字确定得到，如果从所述目标网页的各位置区域中确定出某位置区域中包括的数据信息被爬取的关键字最多，则确定该位置区域为第一位置区域。例如，假设所述目标网页中包括评论位置区域、提问位置区域、图像显示位置区域、信息显示位置区域，如果获取到历史记录中记录了评论位置区域的数据信息被爬取的关键字数量为 n，提问位置区域的数据信息被爬取的关键字数量为 m，图像显示位置区域的数据信息被爬取的关键字数量为 x，信息显示位置区域的数据信息被爬取的关键字数量为 y，将各位置区域的数据信息被爬取的关键字数量进行对比，如果 n>m>x>y，则可以确定所述评论位置区域的数据信息被爬取的关键字数量最多，因此确定所述目标网页中的评论位置区域为第一位置区域。

通过该实施方式，在检测到访问该目标网页的终端为爬虫终端时，可以确定该爬虫终端最大概率会爬取该目标网页的第一位置区域的数据信息，从而提高该爬虫终端打开该第一位置区域的至少一个 JS 挖矿脚本的链接的效率，以提高反爬的效率。

在一个实施例中，该终端可以根据获取到的第一 JS 挖矿脚本生成第一挖矿脚本的链接，以及根据获取到的第二 JS 挖矿脚本生成第二挖矿脚本的链接，并将该第一 JS 挖矿脚本的链接嵌入目标网页的第二位置区域，以及将第二 JS 挖矿脚本的链接嵌入目标网页的第三位置区域。

其中，所述第二位置区域是根据历史记录中记录的所述目标网页中被支持所述 WebAssembly 语言的浏览器爬取各位置区域的数据信息的概率确定的，如果从所述目标网页的各位置区域中确定出某位置区域的数据信息被支持所述 WebAssembly 语言的浏览器爬取的概率最大，则确定所述被支持所述 WebAssembly 语言的浏览器爬取的概率最大的位置区域为第二位置区域。

例如，假设所述目标网页中包括评论位置区域、提问位置区域、图像显示位置区域、信息显示位置区域，如果获取到历史记录中记录了评论位置区域和提问位置区域的数据信息被支持所述 WebAssembly 语言的浏览器爬取，且所述评论位置区域的数据信息被支持所述 WebAssembly 语言的浏览器爬取的概率为 80%，所述提问位置区域的数据信息被支持所述 WebAssembly 语言的浏览器爬取的概率为 50%，将所述评论位置区域的数据信息被支持所述 WebAssembly 语言的浏览器爬取的概率与所述提问位置区域的数据信息被支持所述 WebAssembly 语言的浏览器爬取的概率进行对比，确定所述评论位置区域的数据信息被支持所述 WebAssembly 语言的浏览器爬取的概率 80%最大，因此确定所述目标网页中的评

论位置区域为所述第二位置区域。

在一个实施例中，所述第二位置区域可以是根据历史记录中记录的所述目标网页中各位置区域的数据信息被支持所述 WebAssembly 语言的浏览器爬取的关键字确定的，如果从所述目标网页的各位置区域中确定出某位置区域中包括的数据信息被支持所述 WebAssembly 语言的浏览器爬取的关键字最多，则确定该位置区域为第二位置区域。

例如，假设所述目标网页中包括评论位置区域、提问位置区域、图像显示位置区域、信息显示位置区域，如果获取到历史记录中记录了评论位置区域的数据信息被支持所述 WebAssembly 语言的浏览器爬取的关键字数量为 n，提问位置区域的数据信息被支持所述 WebAssembly 语言的浏览器爬取的关键字数量为 m，图像显示位置区域的数据信息被支持所述 WebAssembly 语言的浏览器爬取的关键字数量为 x，信息显示位置区域的数据信息被支持所述 WebAssembly 语言的浏览器爬取的关键字数量为 y，将各位置区域的数据信息被支持所述 WebAssembly 语言的浏览器爬取的关键字数量进行对比，如果 n>m>x>y，则可以确定所述评论位置区域的数据信息被支持所述 WebAssembly 语言的浏览器爬取的关键字数量最多，因此确定所述目标网页中的评论位置区域为第二位置区域。

在一个实施例中，所述第三位置区域是根据历史记录中记录的所述目标网页中被不支持所述 WebAssembly 语言的浏览器爬取各位置区域的数据信息的概率确定的，如果从所述目标网页的各位置区域中确定出某位置区域的数据信息被不支持所述 WebAssembly 语言的浏览器爬取的概率最大，则确定所述被不支持所述 WebAssembly 语言的浏览器爬取的概率最大的位置区域为第三位置区域。

例如，假设所述目标网页中包括评论位置区域、提问位置区域、图像显示位置区域、信息显示位置区域，如果获取到历史记录中记录了图像显示位置区域和信息显示位置区域的数据信息被不支持所述 WebAssembly 语言的浏览器爬取，且所述图像显示位置区域的数据信息被不支持所述 WebAssembly 语言的浏览器爬取的概率为 60%，所述信息显示位置区域的数据信息被不支持所述 WebAssembly 语言的浏览器爬取的概率为 90%，将所述图像显示位置区域的数据信息被不支持所述 WebAssembly 语言的浏览器爬取的概率与所述信息显示位置区域的数据信息被支持所述 WebAssembly 语言的浏览器爬取的概率进行对比，确定所述信息显示位置区域的数据信息被不支持所述 WebAssembly 语言的浏览器爬取的概率 90%最大，因此确定所述目标网页中的信息显示位置区域为所述第三位置区域。

在一个实施例中，所述第三位置区域还可以是根据历史记录中记录的所述目标网页中各位置区域的数据信息被不支持所述 WebAssembly 语言的浏览器爬取的关键字确定的，如果从所述目标网页的各位置区域中确定出某位置区域中包括的数据信息被不支持所述 WebAssembly 语言的浏览器爬取的关键字最多，则确定该位置区域为第三位置区域。

例如，假设所述目标网页中包括评论位置区域、提问位置区域、图像显示位置区域、信息显示位置区域，如果获取到历史记录中记录了评论位置区域的数据信息被不支持所述 WebAssembly 语言的浏览器爬取的关键字数量为 n，提问位置区域的数据信息被不支持所述 WebAssembly 语言的浏览器爬取的关键字数量为 m，图像显示位置区域的数据信息被不支持所述 WebAssembly 语言的浏览器爬取的关键字数量为 x，信息显示位置区域的数据信息被不支持所述 WebAssembly 语言的浏览器爬取的关键字数量为 y，将各位置区域的数据信息被支持所述 WebAssembly 语言的浏览器爬取的关键字数量进行对比，如果 n<m<x<y，则可以确定所述信息显示位置区域的数据信息被不支持所述 WebAssembly 语言的浏览器爬取的关键字数量最多，因此确定所述目标网页中的信息显示位置区域为所述第三位置区域。

通过该实施方式，在检测到访问该目标网页的终端为爬虫终端时，如果访问该目标网页的终端使用的浏览器支持 WebAssembly，则可以确定该爬虫终端最大概率会先爬取该目标网页的第二位置区域的数据信息。如果访问该目标网页的终端使用的浏览器不支持 WebAssembly，则可以确定该爬虫终端最大概率会先爬取该目标网页的第三位置区域的数据信息。因此根据访问该目标网页的终端使用的浏览器所支持的语言，访问该目标网页的终端最大概率会加载并执行与该爬虫终端使用的浏览器所支持的语言对应的 JS 挖矿脚本的链接，从而进一步提高了反爬的效率和有效性。

在一个实施例中，该终端在将至少一个 JS 挖矿脚本的链接嵌入目标网页时，可以在该至少一个 JS 挖矿脚本的链接上添加提示信息，以提醒访问该目标网页的正常用户不要点击此链接，并将添加提示信息的该至少一个 JS 挖矿脚本的链接嵌入该目标网页的任意位置区域。

例如，如果该终端将至少一个 JS 挖矿脚本的链接嵌入目标网页，则可以在至少一个 JS 挖矿脚本的链接上添加“请勿点击”的提示信息，然后将添加该提示信息的至少一个 JS 挖矿脚本的链接嵌入到该目标网页上的第一位置区域，其中所述目标位位置区域的解释如前所述，此处不再赘述。

又例如，如果该终端分别将第一 JS 挖矿脚本的链接嵌入目标网页的第二位置区域，以及将第二 JS 挖矿脚本的链接嵌入目标网页的第三位置区域，则可以在该第一 JS 挖矿脚本的链接和该第二 JS 挖矿脚本的链接上添加“请勿点击”的提示信息，然后分别将添加该提示信息的第一 JS 挖矿脚本的链接嵌入到该目标网页上的第二位置区域，以及将添加该提示信息的第二 JS 挖矿脚本的链接嵌入到该目标网页上的第三位置区域，其中，所述第二位置区域和第三位置区域的解释如前所述，此处不再赘述。

通过该实施方式可以在一定程度上避免访问该目标网页的正常终端的用户不小心点击到嵌入该目标网页的所述至少一个 JS 挖矿脚本的链接，从而避免对访问该目标网页的正常终端造成损害。

S103：当检测到至少一个 JS 挖矿脚本的链接的启动指令时，触发访问目标网页的终端加载并执行所述至少一个 JS 挖矿脚本。

本申请实施例中，在检测到有终端访问该目标网页，如果检测到嵌入该目标网页的至少一个 JS 挖矿脚本的链接被打开，则确定访问该目标网页的终端为爬虫终端，并触发访问所述目标网页的终端加载并执行所述至少一个 JS 挖矿脚本，以大量消耗访问所述目标网页的终端的 CPU 资源，到达反爬的目的，并防止了该终端绕过 JS 渲染爬取该目标网页的数据信息，从而提高反爬的有效性以及增强了互联网的安全。

本申请实施例中，终端可以通过将至少一个 JS 挖矿脚本的链接嵌入目标网页，当检测到所述至少一个 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端加载并执行至少一个 JS 挖矿脚本，从而防止该访问所述目标网页的终端绕过 JS 渲染爬取数据，提高了反爬的有效性。

参见图 2，图 2 是本申请实施例提供的另一种反爬虫方法的示意流程图，如图 2 所示，该方法可以由终端执行，该终端的具体解释如前所述，此处不再赘述。具体地，本申请实施例的所述方法包括如下步骤。

S201：获取 JS 挖矿脚本库。

本申请实施例中，终端可以建立 JS 挖矿脚本库，该 JS 挖矿脚本库中包括第一 JS 挖矿脚本和第二 JS 挖矿脚本。其中，所述第一 JS 挖矿脚本和第二 JS 挖矿脚本的解释如前所述，此处不再赘述。

S202：生成至少一个 JS 挖矿脚本的链接，将至少一个 JS 挖矿脚本的链接嵌入目标网页的第一位置区域。

本申请实施例中，终端可以根据获取到的 JS 挖矿脚本库生成该至少一个 JS 挖矿脚本的链接，并将该至少一个 JS 挖矿脚本的链接嵌入到目标网页的标位置区域。其中，所述第一位置区域的具体如前所述，此处不再赘述。

S203：当检测到至少一个 JS 挖矿脚本的链接的启动指令时，触发访问该目标网页的终端根据访问该目标网页的终端使用的浏览器所支持的语言，选取第一 JS 挖矿脚本或第二 JS 挖矿脚本。

本申请实施例中，终端可以在检测到所述至少一个 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端根据访问该目标网页的终端使用的浏览器所支持的语言，选取所述第一 JS 挖矿脚本或第二 JS 挖矿脚本。

其中，访问目标网页的终端可以根据访问该目标网页的终端所使用的浏览的版本号来判断访问该目标网页的终端所使用的浏览器是否支持 WebAssembly 语言语言。在一个实施例中，在检测到所述至少一个 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端判断访问该目标网页的终端使用的浏览器版本号是否为预设版本号集合中的任意一种。如果判断出访问该目标网页的终端使用的浏览器版本号为预设版本号集合中的任意一种，则确定访问该目标网页的终端使用的浏览器支持 WebAssembly

语言，因此从该 JS 挖矿脚本库中选取第一 JS 挖矿脚本。如果判断出访问该目标网页的终端使用的浏览器版本号不是预设版本号集合中的任意一种，则确定访问该目标网页的终端使用的浏览器不支持 WebAssembly 语言，从而从该 JS 挖矿脚本库中选取第二 JS 挖矿脚本。

例如，假设预设版本号集合为{ Firefox1.0、Firefox9、Firefox8、Firefox7、Firefox6、Firefox5、Firefox4、Firefox3.6、Firefox3、IE9、IE8、IE7}，当检测到至少一个 JS 挖矿脚本的链接的启动指令时，触发访问目标网页的终端判断该访问目标网页的终端所使用的浏览器版本号是否为预设版本号集合中的任意一种，如果判断出访问目标网页的终端所使用的浏览器版本号为 IE8，则可以确定该访问目标网页的终端使用的浏览器为预设版本号集合中的一种，因此该访问目标网页的终端使用的浏览器支持 WebAssembly 语言，从该 JS 挖矿脚本库中选取第一 JS 挖矿脚本。

又例如，假设预设版本号集合为{ Firefox1.0、Firefox9、Firefox8、Firefox7、Firefox6、Firefox5、Firefox4、Firefox3.6、Firefox3、IE9、IE8、IE7}，当检测到至少一个 JS 挖矿脚本的链接的启动指令时，触发访问目标网页的终端判断该访问目标网页的终端所使用的浏览器版本号是否为预设版本号集合中的任意一种，如果判断出访问目标网页的终端所使用的浏览器版本号为 IE5，则可以确定该访问目标网页的终端使用的浏览器不是预设版本号集合中的任意一种，因此该访问目标网页的终端使用的浏览器不支持 WebAssembly 语言，从该 JS 挖矿脚本库中选取第一 JS 挖矿脚本。

S204：调用访问该目标网页的终端使用的浏览器加载并执行第一 JS 挖矿脚本或第二 JS 挖矿脚本。

本申请实施例中，终端可以调用访问目标网页的终端使用的浏览器加载并执行访问所述目标网页的终端速选取的第一 JS 挖矿脚本或第二 JS 挖矿脚本。

在一个实施例中，如果检测到访问所述目标网页的终端从 JS 挖矿脚本库中选取的脚本是第一 JS 挖矿脚本，则终端可以调用访问所述目标网页的终端所使用的浏览器加载并执行该第一 JS 挖矿脚本。例如，如果检测到访问所述目标网页的终端从 JS 挖矿脚本库中选取的脚本是第一 JS 挖矿脚本，假设确定出访问所述目标网页的终端所使用的浏览器为 IE8，则可以调用访问该目标网页的终端所使用的浏览器 IE8 加载并执行该第一 JS 挖矿脚本。

在一个实施例中，如果检测到访问目标网页的终端从 JS 挖矿脚本库中选取的脚本是第二 JS 挖矿脚本，则终端可以调用访问所述目标网页的终端所使用的不支持 WebAssembly 语言的浏览器加载并执行该第二 JS 挖矿脚本。例如，如果检测到访问目标网页的终端从 JS 挖矿脚本库中选取的脚本是第二 JS 挖矿脚本，假设确定出访问所述目标网页的终端所使用的浏览器为 IE5，则可以调用访问该目标网页的终端所使用的浏览器 IE5 加载并执行该第二 JS 挖矿脚本。

本申请实施例中，终端可以在检测到至少一个 JS 挖矿脚本的链接的启动指令时，触发访问该目标

网页的终端根据访问该目标网页的终端所使用的浏览器支持的语言，选取第一 JS 挖矿脚本或第二 JS 挖矿脚本，并调用访问目标网页的终端使用的浏览器加载并执行访问该目标网页的终端选取的第一 JS 挖矿脚本或第二 JS 挖矿脚本，以消耗访问该目标网页的终端的 CPU 资源，达到反爬的目的，提高反爬的有效性。

参见图 3，图 3 是本申请实施例提供的又一种反爬虫方法的示意流程图，如图 3 所示，该方法可以由终端执行，该终端的具体解释如前所述，此处不再赘述。具体地，本申请实施例的所述方法包括如下步骤。

S301：获取 JS 挖矿脚本库，该 JS 挖矿脚本库包括第一 JS 挖矿脚本和第二 JS 挖矿脚本。

本申请实施例中，终端可以获取 JS 挖矿脚本库，所述 JS 挖矿脚本库包括第一 JS 挖矿脚本和第二 JS 挖矿脚本，所述第一 JS 挖矿脚本和第二 JS 挖矿脚本的解释如前所述，此处不再赘述。

S302：生成第一 JS 挖矿脚本的链接和第二 JS 挖矿脚本的链接。

本申请实施例中，终端可以根据该 JS 挖矿脚本库中包括的第一 JS 挖矿脚本和第二 JS 挖矿脚本，生成第一 JS 挖矿脚本的链接和第二 JS 挖矿脚本的链接。

S303：将该第一 JS 挖矿脚本的链接嵌入该目标网页的第二位置区域，以及将该第二 JS 挖矿脚本的链接嵌入该目标网页的第三位置区域。

本申请实施例中，终端可以将所述第一 JS 挖矿脚本的链接嵌入所述目标网页的所述第二位置区域，以及将所述第二 JS 挖矿脚本的链接嵌入所述目标网页的所述第三位置区域。其中，所述第二位置区域和第三位置区域的解释如前所述，此处不再赘述。例如，假设第二位置区域为目标网页中的评论位置区域，第三位置区域为目标网页中的图像显示位置区域，则终端可以将所述第一 JS 挖矿脚本的链接嵌入所述目标网页的评论位置区域，以及将所述第二 JS 挖矿脚本的链接嵌入所述目标网页的图像显示位置区域。

S304：当检测到该第一 JS 挖矿脚本的链接的启动指令时，触发访问该目标网页的终端判断访问该目标网页的终端使用的浏览器版本号是否为预设版本号集合中的任意一种，如果判断结果为是，则执行步骤 S305，如果判断结果为否，则执行步骤 S306。

本申请实施例中，当检测到所述第一 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端判断访问所述目标网页的终端所使用的浏览器版本号是否为预设版本号集合中的任意一种，如果判断出访问该目标网页的终端使用的浏览器版本号为预设版本号集合中的任意一种，则执行步骤 S305。如果判断出访问该目标网页的终端使用的浏览器版本号不是预设版本号集合中的任意一种，则执行步骤 S306。其中，触发访问所述目标网页的终端判断访问所述目标网页的终端所使用的浏览器版本号是否为

预设版本号集合中的任意一种的具体实施过程和举例如前所述，此处不再赘述。

例如，当检测到嵌入到目标网页中评论位置区域的第一 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端判断访问所述目标网页的终端所使用的浏览器版本号是否为预设版本号集合中的任意一种。又例如，当检测到嵌入到目标网页中图像位置区域的第二 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端判断访问所述目标网页的终端所使用的浏览器版本号是否为预设版本号集合中的任意一种。

S305：确定访问该目标网页的终端使用的浏览器支持 WebAssembly 语言，并触发访问该目标网页的终端使用的浏览器加载并执行该第一 JS 挖矿脚本。

本申请实施例中，如果判断出访问该目标网页的终端使用的浏览器版本号为预设版本号集合中的任意一种，则确定访问所述目标网页的终端使用的浏览器支持所述 WebAssembly 语言，并触发访问所述目标网页的终端使用的浏览器加载并执行所述第一 JS 挖矿脚本。其中，触发访问所述目标网页的终端使用的浏览器加载并执行所述第一 JS 挖矿脚本的具体实施过程以及举例说明如前所述，此处不再赘述。

S306：在检测到第二 JS 挖矿脚本的链接的启动指令时，触发访问该目标网页的终端使用的浏览器加载并执行该第二 JS 挖矿脚本。

本申请实施例中，如果判断出该终端使用的浏览器版本号不是预设版本号集合中的任意一种，则在检测到所述第二 JS 挖矿脚本的链接的启动指令时，确定访问所述目标网页的终端使用的浏览器不支持所述 WebAssembly 语言，并触发访问所述目标网页的终端使用的浏览器加载并执行所述第二 JS 挖矿脚本。其中，触发访问所述目标网页的终端使用的浏览器加载并执行所述第二 JS 挖矿脚本的具体实施过程以及举例说明如前所述，此处不再赘述。

本申请实施例中，终端可以将生成的第一 JS 挖矿脚本的链接嵌入目标网页的第二位置区域，以及将生成的第二 JS 挖矿脚本的链接嵌入目标网页的第三位置区域，当检测到第一 JS 挖矿脚本的链接的启动指令时，根据访问所述目标网页的终端所使用的浏览器版本号，调用访问所述目标网页的终端所使用的浏览器加载并执行第一 JS 挖矿脚本或第二 JS 挖矿脚本。通过这种方式可以大量消耗访问目标网页的终端的 CPU 资源，达到反爬的目的，进一步提高了反爬的效率和有效性，从而增强了互联网的安全性。

本申请实施例还提供了一种终端，该终端用于执行前述任一项所述的方法的单元。具体地，参见图 4，图 4 是本申请实施例提供的一种终端的示意框图。本实施例的终端包括：获取单元 401、嵌入单元 402、触发单元 403。

获取单元 401，用于获取 JS 挖矿脚本库，所述 JS 挖矿脚本库包括至少一个 JS 挖矿脚本。

进一步地，获取单元 401，具体用于建立所述 JS 挖矿脚本库，所述 JS 挖矿脚本库包括第一 JS 挖矿

脚本和第二 JS 挖矿脚本；其中，所述第一 JS 挖矿脚本是 WebAssembly 语言的脚本，所述第二 JS 挖矿脚本是 JS 语言的脚本。

嵌入单元 402，用于生成所述至少一个 JS 挖矿脚本的链接，将所述至少一个 JS 挖矿脚本的链接嵌入目标网页。

进一步地，嵌入单元 402，具体用于根据历史记录中所述目标网页中各位置区域的数据信息被爬取的概率确定第一位置区域，所述第一位置区域是所述目标网页中各位置区域的数据信息被爬取概率最大的位置区域；将所述至少一个 JS 挖矿脚本的链接嵌入所述目标网页的所述第一位置区域。

进一步地，嵌入单元 402，具体用于生成所述第一 JS 挖矿脚本的链接和所述第二 JS 挖矿脚本的链接；根据历史记录中所述目标网页中被支持所述 WebAssembly 语言的浏览器爬取各位置区域的数据信息的概率确定第二位置区域，所述第二位置区域是所述目标网页中各位置区域的数据信息被支持所述 WebAssembly 语言的浏览器爬取概率最大的位置区域；根据历史记录中所述目标网页中被不支持所述 WebAssembly 语言的浏览器爬取各位置区域的数据信息的概率确定第三位置区域，所述第三位置区域是所述目标网页中各位置区域的数据信息被不支持所述 WebAssembly 语言的浏览器爬取概率最大的位置区域；将所述第一 JS 挖矿脚本的链接嵌入所述目标网页的所述第二位置区域，以及将所述第二 JS 挖矿脚本的链接嵌入所述目标网页的所述第三位置区域。

触发单元 403，用于当检测到所述至少一个 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端加载并执行所述至少一个 JS 挖矿脚本。

进一步地，触发单元 403，用于当检测到所述至少一个 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端根据所述终端使用的浏览器所支持的语言，选取所述第一 JS 挖矿脚本或第二 JS 挖矿脚本；调用所述终端使用的浏览器加载并执行所述终端选取的所述第一 JS 挖矿脚本或第二 JS 挖矿脚本。

进一步地，触发单元 403，用于触发访问所述目标网页的终端判断所述终端使用的浏览器版本号是否为预设版本号集合中的任意一种；如果判断结果为是，则确定所述终端使用的浏览器支持 WebAssembly 语言，从所述 JS 挖矿脚本库中选取所述第一 JS 挖矿脚本；如果判断结果为否，则确定所述终端使用的浏览器不支持 WebAssembly 语言，从所述 JS 挖矿脚本库中选取所述第二 JS 挖矿脚本。

进一步地，触发单元 403，用于当检测到所述第一 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端判断所述终端使用的浏览器版本号是否为预设版本号集合中的任意一种；如果判断结果为是，则确定所述终端使用的浏览器支持所述 WebAssembly 语言，并触发访问所述目标网页的终端使用的浏览器加载并执行所述第一 JS 挖矿脚本；如果判断结果为否，则在检测到所述第二 JS 挖矿脚本的

链接的启动指令时，确定所述终端使用的浏览器不支持所述 WebAssembly 语言，并触发访问所述目标网页的终端使用的浏览器加载并执行所述第二 JS 挖矿脚本。

本申请实施例中，终端的获取单元 401 可以通过生成所述至少一个 JS 挖矿脚本的链接，嵌入单元 402 将所述至少一个 JS 挖矿脚本的链接嵌入目标网页，触发单元 403 当检测到所述至少一个 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端加载并执行所述至少一个 JS 挖矿脚本，从而防止爬虫终端绕过 JS 渲染爬取数据，提高了反爬的有效性。

参见图 5，图 5 是本申请实施例提供的另一种终端示意框图。如图所示的本实施例中的终端可以包括：一个或多个处理器 501；一个或多个输入设备 502，一个或多个输出设备 503 和存储器 504。上述处理器 501、输入设备 402、输出设备 503 和存储器 504 通过总线 505 连接。存储器 504 用于存储计算机程序，所述计算机程序包括程序指令，处理器 501 用于执行存储器 504 存储的程序指令。其中，处理器 501 被配置用于调用所述程序指令执行 ：

获取 JS 挖矿脚本库，所述 JS 挖矿脚本库包括至少一个 JS 挖矿脚本；

生成所述至少一个 JS 挖矿脚本的链接，将所述至少一个 JS 挖矿脚本的链接嵌入目标网页；

当检测到所述至少一个 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端加载并执行所述至少一个 JS 挖矿脚本。

进一步地，所述处理器 501 用于执行如下步骤：

建立所述 JS 挖矿脚本库，所述 JS 挖矿脚本库包括第一 JS 挖矿脚本和第二 JS 挖矿脚本；

其中，所述第一 JS 挖矿脚本是 WebAssembly 语言的脚本，所述第二 JS 挖矿脚本是 JS 语言的脚本。

进一步地，所述处理器 501 用于执行如下步骤：

根据历史记录中所述目标网页中各位置区域的数据信息被爬取的概率确定第一位置区域，所述第一位置区域是所述目标网页中各位置区域的数据信息被爬取概率最大的位置区域；

将所述至少一个 JS 挖矿脚本的链接嵌入所述目标网页的所述第一位置区域。

进一步地，所述处理器 501 用于执行如下步骤：

当检测到所述至少一个 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端根据所述终端使用的浏览器所支持的语言，选取所述第一 JS 挖矿脚本或第二 JS 挖矿脚本；

调用所述终端使用的浏览器加载并执行所述终端选取的所述第一 JS 挖矿脚本或第二 JS 挖矿脚本。

进一步地，所述处理器 501 用于执行如下步骤：

触发访问所述目标网页的终端判断所述终端使用的浏览器版本号是否为预设版本号集合中的任意一种；

如果判断结果为是，则确定所述终端使用的浏览器支持 WebAssembly 语言，从所述 JS 挖矿脚本库中选取所述第一 JS 挖矿脚本;

如果判断结果为否，则确定所述终端使用的浏览器不支持 WebAssembly 语言，从所述 JS 挖矿脚本库中选取所述第二 JS 挖矿脚本。

进一步地，所述处理器 501 用于执行如下步骤:

生成所述第一 JS 挖矿脚本的链接和所述第二 JS 挖矿脚本的链接;

根据历史记录中所述目标网页中被支持所述 WebAssembly 语言的浏览器爬取各位置区域的数据信息的概率确定第二位置区域，所述第二位置区域是所述目标网页中各位置区域的数据信息被支持所述 WebAssembly 语言的浏览器爬取概率最大的位置区域;

根据历史记录中所述目标网页中被不支持所述 WebAssembly 语言的浏览器爬取各位置区域的数据信息的概率确定第三位置区域，所述第三位置区域是所述目标网页中各位置区域的数据信息被不支持所述 WebAssembly 语言的浏览器爬取概率最大的位置区域;

将所述第一 JS 挖矿脚本的链接嵌入所述目标网页的所述第二位置区域，以及将所述第二 JS 挖矿脚本的链接嵌入所述目标网页的所述第三位置区域。

进一步地，所述处理器 501 用于执行如下步骤:

当检测到所述第一 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端判断所述终端使用的浏览器版本号是否为预设版本号集合中的任意一种;

如果判断结果为是，则确定所述终端使用的浏览器支持所述 WebAssembly 语言，并触发访问所述目标网页的终端使用的浏览器加载并执行所述第一 JS 挖矿脚本;

如果判断结果为否，则在检测到所述第二 JS 挖矿脚本的链接的启动指令时，确定所述终端使用的浏览器不支持所述 WebAssembly 语言，并触发访问所述目标网页的终端使用的浏览器加载并执行所述第二 JS 挖矿脚本。

本申请实施例中，终端可以在检测到所述至少一个 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端加载并执行所述至少一个 JS 挖矿脚本，从而大量消耗爬虫终端的 CPU 资源，使爬虫终端无法正常使用，以达到反爬的目的，以及可以防止爬虫终端绕过 JS 渲染爬取数据，提高了反爬的有效性。

具体实现中，本申请实施例中所描述的处理器 501、输入设备 502、输出设备 503 可执行本申请实施例提供的 反爬虫方法的图 1、图 2 或图 3 所述的方法实施例中所描述的实现方式，也可执行本申请实施例图 4 所描述的终端的实现方式，在此不再赘述。

本申请实施例中还提供了一种计算机可读存储介质，所述计算机可读存储介质存储有计算机程序，所述计算机程序被处理器执行时实现图 1、图 2 或图 3 所对应实施例中描述的反爬虫方法，也可实现本申请图 4 或图 5 所对应实施例的终端，在此不再赘述。

所述计算机可读存储介质可以是前述任一实施例所述的终端的内部存储单元，例如终端的硬盘或内存。所述计算机可读存储介质也可以是所述终端的外部存储设备，例如所述终端上配备的插接式硬盘，智能存储卡（Smart Media Card, SMC），安全数字（Secure Digital, SD）卡，闪存卡（Flash Card）等。进一步地，所述计算机可读存储介质还可以既包括所述终端的内部存储单元也包括外部存储设备。所述计算机可读存储介质用于存储所述计算机程序以及所述终端所需的其他程序和数据。所述计算机可读存储介质还可以用于暂时地存储已经输出或者将要输出的数据。

以上所述，仅为本申请的部分实施方式，但本申请的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本申请揭露的技术范围内，可轻易想到各种等效的修改或替换，这些修改或替换都应涵盖在本申请的保护范围之内。

权 利 要 求

1.一种反爬虫方法，其特征在于，包括:

获取 JS 挖矿脚本库，所述 JS 挖矿脚本库包括至少一个 JS 挖矿脚本;

生成所述至少一个 JS 挖矿脚本的链接，将所述至少一个 JS 挖矿脚本的链接嵌入目标网页;

当检测到所述至少一个 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端加载并执行所述至少一个 JS 挖矿脚本。

2.根据权利要求 1 所述的方法，其特征在于，所述获取 JS 挖矿脚本库，包括:

建立所述 JS 挖矿脚本库，所述 JS 挖矿脚本库包括第一 JS 挖矿脚本和第二 JS 挖矿脚本;

其中，所述第一 JS 挖矿脚本是 WebAssembly 语言的脚本，所述第二 JS 挖矿脚本是 JS 语言的脚本。

3.根据权利要求 1 所述的方法，其特征在于，所述将所述至少一个 JS 挖矿脚本的链接嵌入目标网页，包括:

根据历史记录中所述目标网页中各位置区域的数据信息被爬取的概率确定第一位置区域，所述第一位置区域是所述目标网页中各位置区域的数据信息被爬取概率最大的位置区域;

将所述至少一个 JS 挖矿脚本的链接嵌入所述目标网页的所述第一位置区域。

4.根据权利要求 2 所述的方法，其特征在于，所述当检测到所述至少一个 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端加载并执行所述至少一个 JS 挖矿脚本，包括:

当检测到所述至少一个 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端根据所述终端使用的浏览器所支持的语言，选取所述第一 JS 挖矿脚本或第二 JS 挖矿脚本;

调用所述终端使用的浏览器加载并执行所述终端选取的所述第一 JS 挖矿脚本或第二 JS 挖矿脚本。

5.根据权利要求 4 所述的方法，其特征在于，所述触发访问所述目标网页的终端根据所述终端使用的浏览器所支持的语言，选取所述第一 JS 挖矿脚本或第二 JS 挖矿脚本，包括:

触发访问所述目标网页的终端判断所述终端使用的浏览器版本号是否为预设版本号集合中的任意一种;

如果判断结果为是，则确定所述终端使用的浏览器支持 WebAssembly 语言，从所述 JS 挖矿脚本库中选取所述第一 JS 挖矿脚本;

如果判断结果为否，则确定所述终端使用的浏览器不支持 WebAssembly 语言，从所述 JS 挖矿脚本库中选取所述第二 JS 挖矿脚本。

6.根据权利要求 2 所述的方法，其特征在于，所述将所述至少一个 JS 挖矿脚本的链接嵌入目标网页，

包括:

生成所述第一 JS 挖矿脚本的链接和所述第二 JS 挖矿脚本的链接;

根据历史记录中所述目标网页中被支持所述 WebAssembly 语言的浏览器爬取各位置区域的数据信息的概率确定第二位置区域，所述第二位置区域是所述目标网页中各位置区域的数据信息被支持所述 WebAssembly 语言的浏览器爬取概率最大的位置区域;

根据历史记录中所述目标网页中被不支持所述 WebAssembly 语言的浏览器爬取各位置区域的数据信息的概率确定第三位置区域，所述第三位置区域是所述目标网页中各位置区域的数据信息被不支持所述 WebAssembly 语言的浏览器爬取概率最大的位置区域;

将所述第一 JS 挖矿脚本的链接嵌入所述目标网页的所述第二位置区域，以及将所述第二 JS 挖矿脚本的链接嵌入所述目标网页的所述第三位置区域。

7.根据权利要求 6 所述的方法，其特征在于，所述当检测到所述至少一个 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端加载并执行所述至少一个 JS 挖矿脚本，包括:

当检测到所述第一 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端判断所述终端使用的浏览器版本号是否为预设版本号集合中的任意一种;

如果判断结果为是，则确定所述终端使用的浏览器支持所述 WebAssembly 语言，并触发访问所述目标网页的终端使用的浏览器加载并执行所述第一 JS 挖矿脚本;

如果判断结果为否，则在检测到所述第二 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端使用的浏览器加载并执行所述第二 JS 挖矿脚本。

8.一种终端，其特征在于，包括:

获取单元，用于获取 JS 挖矿脚本库，所述 JS 挖矿脚本库包括至少一个 JS 挖矿脚本;

嵌入单元，用于生成所述至少一个 JS 挖矿脚本的链接，将所述至少一个 JS 挖矿脚本的链接嵌入目标网页;

触发单元，用于当检测到所述至少一个 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端加载并执行所述至少一个 JS 挖矿脚本。

9. 根据权利要求 8 所述的终端，其特征在于，所述获取单元 401 获取 JS 挖矿脚本库时，具体用于:

建立所述 JS 挖矿脚本库，所述 JS 挖矿脚本库包括第一 JS 挖矿脚本和第二 JS 挖矿脚本;

其中，所述第一 JS 挖矿脚本是 WebAssembly 语言的脚本，所述第二 JS 挖矿脚本是 JS 语言的脚本。

10. 根据权利要求 8 所述的终端，其特征在于，所述嵌入单元 402 将所述至少一个 JS 挖矿脚本的链接嵌入目标网页时，具体用于:

根据历史记录中所述目标网页中各位置区域的数据信息被爬取的概率确定第一位置区域，所述第一位置区域是所述目标网页中各位置区域的数据信息被爬取概率最大的位置区域；

将所述至少一个 JS 挖矿脚本的链接嵌入所述目标网页的所述第一位置区域。

11、根据权利要求 9 所述的终端，其特征在于，所述触发单元 403 当检测到所述至少一个 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端加载并执行所述至少一个 JS 挖矿脚本时，具体用于：

当检测到所述至少一个 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端根据所述终端使用的浏览器所支持的语言，选取所述第一 JS 挖矿脚本或第二 JS 挖矿脚本；

调用所述终端使用的浏览器加载并执行所述终端选取的所述第一 JS 挖矿脚本或第二 JS 挖矿脚本。

12. 根据权利要求 11 所述的终端，其特征在于，所述触发单元 403 触发访问所述目标网页的终端根据所述终端使用的浏览器所支持的语言，选取所述第一 JS 挖矿脚本或第二 JS 挖矿脚本时，具体用于：

触发访问所述目标网页的终端判断所述终端使用的浏览器版本号是否为预设版本号集合中的任意一种；

如果判断结果为是，则确定所述终端使用的浏览器支持 WebAssembly 语言，从所述 JS 挖矿脚本库中选取所述第一 JS 挖矿脚本；

如果判断结果为否，则确定所述终端使用的浏览器不支持 WebAssembly 语言，从所述 JS 挖矿脚本库中选取所述第二 JS 挖矿脚本。

13. 根据权利要求 9 所述的终端，其特征在于，所述嵌入单元 402 将所述至少一个 JS 挖矿脚本的链接嵌入目标网页时，具体用于：

生成所述第一 JS 挖矿脚本的链接和所述第二 JS 挖矿脚本的链接；

根据历史记录中所述目标网页中被支持所述 WebAssembly 语言的浏览器爬取各位置区域的数据信息的概率确定第二位置区域，所述第二位置区域是所述目标网页中各位置区域的数据信息被支持所述 WebAssembly 语言的浏览器爬取概率最大的位置区域；

根据历史记录中所述目标网页中被不支持所述 WebAssembly 语言的浏览器爬取各位置区域的数据信息的概率确定第三位置区域，所述第三位置区域是所述目标网页中各位置区域的数据信息被不支持所述 WebAssembly 语言的浏览器爬取概率最大的位置区域；

将所述第一 JS 挖矿脚本的链接嵌入所述目标网页的所述第二位置区域，以及将所述第二 JS 挖矿脚本的链接嵌入所述目标网页的所述第三位置区域。

14. 根据权利要求 13 所述的终端，其特征在于，所述触发单元 403 当检测到所述至少一个 JS 挖矿

脚本的链接的启动指令时，触发访问所述目标网页的终端加载并执行所述至少一个 JS 挖矿脚本时，具体用于：

当检测到所述第一 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端判断所述终端使用的浏览器版本号是否为预设版本号集合中的任意一种；

如果判断结果为是，则确定所述终端使用的浏览器支持所述 WebAssembly 语言，并触发访问所述目标网页的终端使用的浏览器加载并执行所述第一 JS 挖矿脚本；

如果判断结果为否，则在检测到所述第二 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端使用的浏览器加载并执行所述第二 JS 挖矿脚本。

15. 一种终端，其特征在于，包括处理器、输入设备、输出设备和存储器，所述处理器、输入设备、输出设备和存储器相互连接，其中，所述存储器用于存储计算机程序，所述计算机程序包括程序指令，所述处理器被配置用于调用所述程序指令以执行如下步骤：

获取 JS 挖矿脚本库，所述 JS 挖矿脚本库包括至少一个 JS 挖矿脚本；

生成所述至少一个 JS 挖矿脚本的链接，将所述至少一个 JS 挖矿脚本的链接嵌入目标网页；

当检测到所述至少一个 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端加载并执行所述至少一个 JS 挖矿脚本。

16. 根据权利要求 15 所述的终端，其特征在于，所述处理器获取 JS 挖矿脚本库时，具体用于：

建立所述 JS 挖矿脚本库，所述 JS 挖矿脚本库包括第一 JS 挖矿脚本和第二 JS 挖矿脚本；

其中，所述第一 JS 挖矿脚本是 WebAssembly 语言的脚本，所述第二 JS 挖矿脚本是 JS 语言的脚本。

17. 根据权利要求 15 所述的终端，其特征在于，所述处理器所述至少一个 JS 挖矿脚本的链接嵌入目标网页时，具体用于：

根据历史记录中所述目标网页中各位置区域的数据信息被爬取的概率确定第一位置区域，所述第一位置区域是所述目标网页中各位置区域的数据信息被爬取概率最大的位置区域；

将所述至少一个 JS 挖矿脚本的链接嵌入所述目标网页的所述第一位置区域。

18. 根据权利要求 16 所述的终端，其特征在于，所述处理器当检测到所述至少一个 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端加载并执行所述至少一个 JS 挖矿脚本时，具体用于：

当检测到所述至少一个 JS 挖矿脚本的链接的启动指令时，触发访问所述目标网页的终端根据所述终端使用的浏览器所支持的语言，选取所述第一 JS 挖矿脚本或第二 JS 挖矿脚本；

调用所述终端使用的浏览器加载并执行所述终端选取的所述第一 JS 挖矿脚本或第二 JS 挖矿脚本。

19. 根据权利要求 18 所述的终端，其特征在于，所述处理器触发访问所述目标网页的终端根据所述

终端使用的浏览器所支持的语言，选取所述第一 JS 挖矿脚本或第二 JS 挖矿脚本时，具体用于：

触发访问所述目标网页的终端判断所述终端使用的浏览器版本号是否为预设版本号集合中的任意一种；

如果判断结果为是，则确定所述终端使用的浏览器支持 WebAssembly 语言，从所述 JS 挖矿脚本库中选取所述第一 JS 挖矿脚本；

如果判断结果为否，则确定所述终端使用的浏览器不支持 WebAssembly 语言，从所述 JS 挖矿脚本库中选取所述第二 JS 挖矿脚本。

20. 一种计算机可读存储介质，其特征在于，所述计算机存储介质存储有计算机程序，所述计算机程序包括程序指令，所述程序指令当被处理器执行时使所述处理器执行如权利要求 1-7 任一项所述的方法。

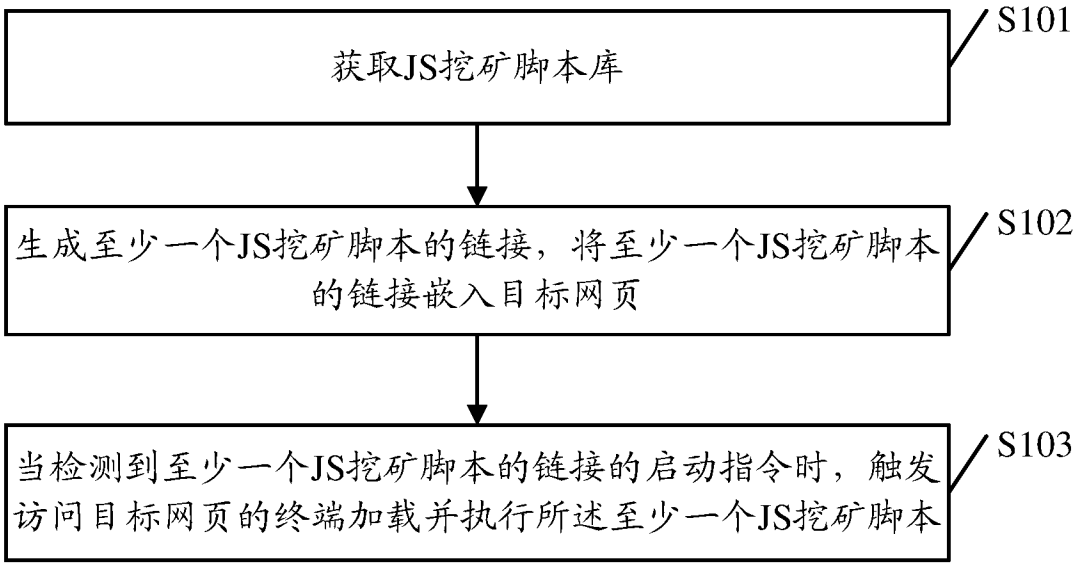

获取JS挖矿脚本库
S101
生成至少一个JS挖矿脚本的链接，将至少一个JS挖矿脚本的链接嵌入目标网页
S102
当检测到至少一个JS挖矿脚本的链接的启动指令时，触发访问目标网页的终端加载并执行所述至少一个JS挖矿脚本
S103

图 1

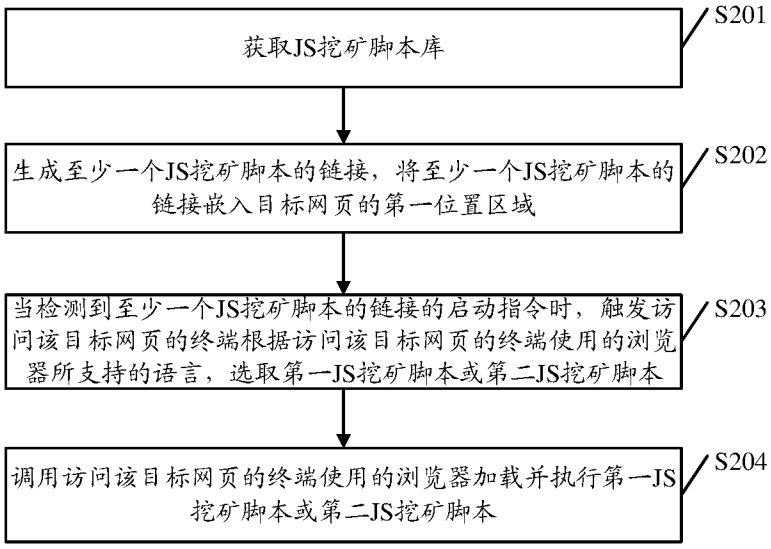

获取JS挖矿脚本库
S201
生成至少一个JS挖矿脚本的链接，将至少一个JS挖矿脚本的链接嵌入目标网页的第一位置区域
S202
当检测到至少一个JS挖矿脚本的链接的启动指令时，触发访问该目标网页的终端根据访问该目标网页的终端使用的浏览器所支持的语言，选取第一JS挖矿脚本或第二JS挖矿脚本
S203
调用访问该目标网页的终端使用的浏览器加载并执行第一JS挖矿脚本或第二JS挖矿脚本
S204

图 2

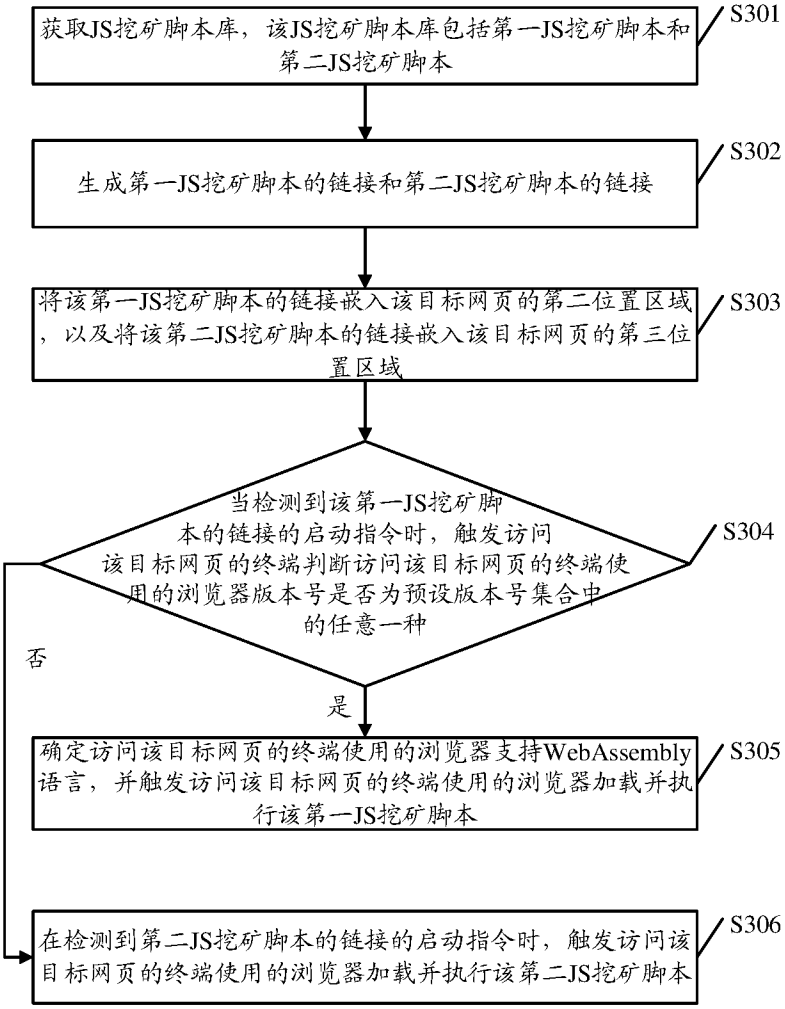

获取JS挖矿脚本库，该JS挖矿脚本库包括第一JS挖矿脚本和第二JS挖矿脚本
S301
生成第一JS挖矿脚本的链接和第二JS挖矿脚本的链接
S302
将该第一JS挖矿脚本的链接嵌入该目标网页的第二位置区域，以及将该第二JS挖矿脚本的链接嵌入该目标网页的第三位置区域
S303
当检测到该第一JS挖矿脚本的链接的启动指令时，触发访问该目标网页的终端判断访问该目标网页的终端使用的浏览器版本号是否为预设版本号集合中的任意一种
S304
否
是
确定访问该目标网页的终端使用的浏览器支持WebAssembly语言，并触发访问该目标网页的终端使用的浏览器加载并执行该第一JS挖矿脚本
S305
在检测到第二JS挖矿脚本的链接的启动指令时，触发访问该目标网页的终端使用的浏览器加载并执行该第二JS挖矿脚本
S306

图 3

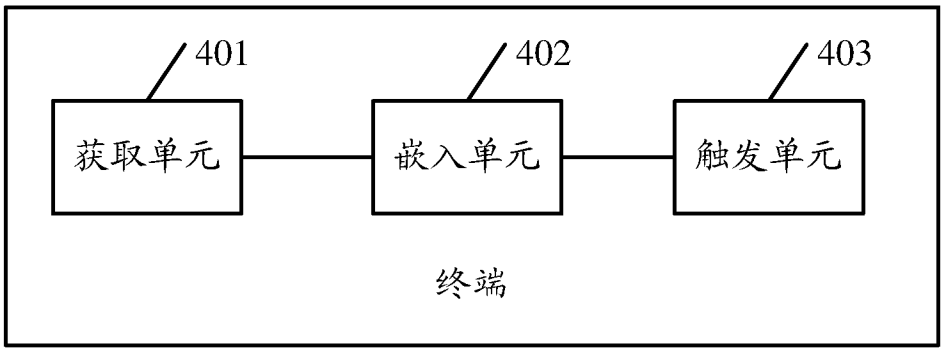

401
402
403
获取单元
嵌入单元
触发单元
终端

图 4

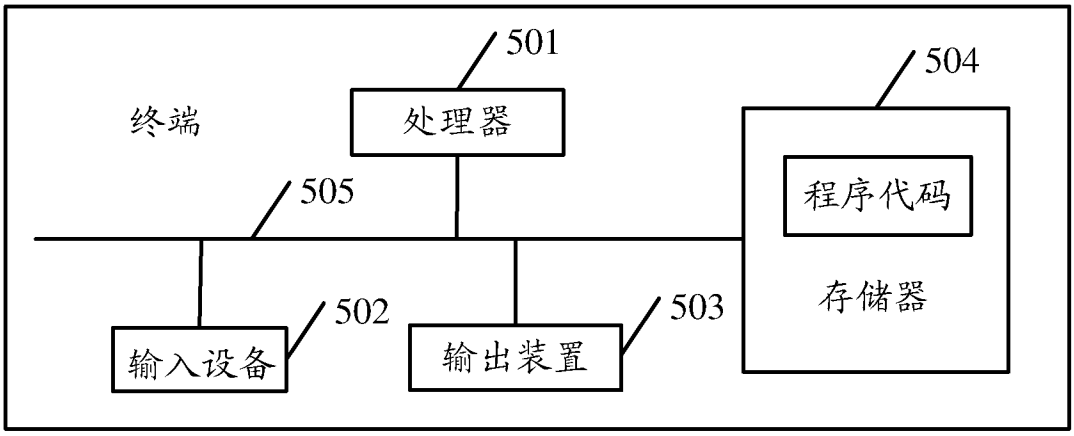

501
504
终端
处理器
程序代码
505
存储器
502
503
输入设备
输出装置

图 5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2018/108672

A. CLASSIFICATION OF SUBJECT MATTER

G06F 21/54(2013.01)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: 反爬虫, 脚本库, 链接, 网页, 终端, JS, crawl, web, assembly, terminal, link

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 105743901 A (CTRIP COMPUTER TECHNOLOGY (SHANGHAI) CO., LTD.) 06 July 2016 (2016-07-06) description, paragraphs [0030]-[0049]	1-20
A	CN 102833212 A (ALIBABA GROUP HOLDING LIMITED) 19 December 2012 (2012-12-19) entire document	1-20
A	CN 107908959 A (BEIJING KNOWNSEC INFORMATION TECHNOLOGY CO., LTD.) 13 April 2018 (2018-04-13) entire document	1-20
A	US 2017346804 A1 (ALTR SOLUTIONS, INC.) 30 November 2017 (2017-11-30) entire document	1-20

☐ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"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)

"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

"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

"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

"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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
14 March 2019	**29 March 2019**

Name and mailing address of the ISA/CN	Authorized officer
China National Intellectual Property Administration (ISA/CN) **No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088** **China**	
Facsimile No. **(86-10)62019451**	Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2018/108672

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 105743901 A	06 July 2016	None	
CN 102833212 A	19 December 2012	HK 1174162 A1	30 September 2016
CN 107908959 A	13 April 2018	None	
US 2017346804 A1	30 November 2017	US 2018302390 A1	18 October 2018
		US 2018077134 A1	15 March 2018

国际检索报告

国际申请号
PCT/CN2018/108672

A. 主题的分类

G06F 21/54(2013.01)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: 反爬虫，脚本库，链接，网页，终端，JS,crawl,web,assembly,terminal,link

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
A	CN 105743901 A (携程计算机技术上海有限公司) 2016年 7月 6日 (2016-07-06) 说明书第 [0030]-[0049]段	1-20
A	CN 102833212 A (阿里巴巴集团控股有限公司) 2012年 12月 19日 (2012-12-19) 全文	1-20
A	CN 107908959 A (北京知道创宇信息技术有限公司) 2018年 4月 13日 (2018-04-13) 全文	1-20
A	US 2017346804 A1 (ALTR SOLUTIONS, INC.) 2017年 11月 30日 (2017-11-30) 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	国际检索报告邮寄日期
2019年 3月 14日	2019年 3月 29日
ISA/CN的名称和邮寄地址 中国国家知识产权局(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 丛磊 电话号码 (86-10)-53961305

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2018/108672

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	105743901	A	2016年 7月 6日		无		
CN	102833212	A	2012年 12月 19日	HK	1174162	A1	2016年 9月 30日
CN	107908959	A	2018年 4月 13日		无		
US	2017346804	A1	2017年 11月 30日	US	2018302390	A1	2018年 10月 18日
				US	2018077134	A1	2018年 3月 15日