(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 5 月 16 日 (16.05.2019)

WIPO | PCT

(10) 国际公布号
WO 2019/091016 A1

(51) 国际专利分类号:
G06F 8/30 (2018.01)

(21) 国际申请号: PCT/CN2018/076961

(22) 国际申请日: 2018 年 2 月 22 日 (22.02.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201711113783.4 2017年11月13日 (13.11.2017) 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八卦岭八卦三路平安大厦六楼, Guangdong 518052 (CN)。

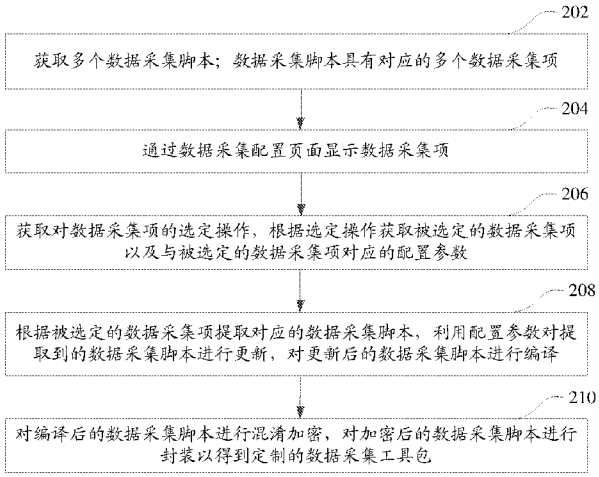
(72) 发明人: 史贵岭 (SHI, Guiling); 中国广东省深圳市福田区八卦岭八卦三路平安大厦六楼, Guangdong 518052 (CN)。 谢培杰 (XIE, Peijie); 中国广东省深圳市福田区八卦岭八卦三路平安大厦六楼, Guangdong 518052 (CN)。 尧德仁 (YAO, Deren); 中国广东省深圳市福田区八卦岭八卦三路平安大厦六楼, Guangdong 518052 (CN)。

(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珠江东路6号4501房（部位：自编01-03和08-12单元）（仅限办公用途）, 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,

(54) **Title:** METHOD AND DEVICE FOR CUSTOMIZING DATA ACQUISITION TOOLKIT, TERMINAL, AND STORAGE MEDIUM

(54) **发明名称:** 数据采集工具包定制方法、装置、终端和存储介质



202 OBTAIN MULTIPLE DATA ACQUISITION SCRIPTS, THE DATA ACQUISITION SCRIPTS HAVING CORRESPONDING MULTIPLE DATA ACQUISITION ITEMS
204 DISPLAY THE DATA ACQUISITION ITEMS BY MEANS OF A DATA ACQUISITION CONFIGURATION PAGE
206 OBTAIN A SELECTION OPERATION ON A DATA ACQUISITION ITEM, AND OBTAIN THE SELECTED DATA ACQUISITION ITEM AND A CONFIGURATION PARAMETER CORRESPONDING TO THE SELECTED DATA ACQUISITION ITEM ACCORDING TO THE SELECTION OPERATION
208 EXTRACT A CORRESPONDING DATA ACQUISITION SCRIPT ACCORDING TO THE SELECTED DATA ACQUISITION ITEM, UPDATE THE EXTRACTED DATA ACQUISITION SCRIPT USING THE CONFIGURATION PARAMETER, AND COMPILE THE UPDATED DATA ACQUISITION SCRIPT
210 PERFORM OBFUSCATION ENCRYPTION ON THE COMPILED DATA ACQUISITION SCRIPT, AND ENCAPSULATE THE ENCRYPTED DATA ACQUISITION SCRIPT TO OBTAIN A CUSTOMIZED DATA ACQUISITION TOOLKIT

图 2

(57) **Abstract:** A method for customizing a data acquisition toolkit, comprising: obtaining multiple data acquisition scripts, the data acquisition scripts having corresponding multiple data acquisition items; displaying the data acquisition items by means of a data acquisition configuration page; obtaining a selection operation on a data acquisition item, and obtaining the selected data acquisition item and a configuration parameter corresponding to the selected data acquisition item according to the selection operation; extracting a corresponding data acquisition script according to the selected data acquisition item, updating the extracted data acquisition script using the configuration parameter, and compiling the updated data acquisition script; and performing obfuscation encryption on the compiled data acquisition script, and encapsulating the encrypted data acquisition script to obtain a customized data acquisition toolkit.

WO 2019/091016 A1

[见续页]

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))。

(57) 摘要：一种数据采集工具包定制方法，包括：获取多个数据采集脚本；数据采集脚本具有对应的多个数据采集项；通过数据采集配置页面显示数据采集项；获取对数据采集项的选定操作，根据选定操作获取被选定的数据采集项以及与被选定的数据采集项对应的配置参数；根据被选定的数据采集项提取对应的数据采集脚本，利用配置参数对提取到的数据采集脚本进行更新，对更新后的数据采集脚本进行编译；对编译后的数据采集脚本进行混淆加密，对加密后的数据采集脚本进行封装以得到定制的数据采集工具包。

数据采集工具包定制方法、装置、终端和存储介质

本申请要求于2017年11月13日提交中国专利局，申请号为2017111137834，申请名称为“数据采集工具包定制方法、装置、终端和存储介质”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

本申请涉及计算机技术领域，特别是涉及数据采集工具包定制方法、装置、终端和存储介质。

背景技术

为了更好的推广运营APP（Application，应用程序），APP开发商需要采集APP中的用户行为数据，并对采集到的数据进行分析。目前，用户行为数据采集方式主要是针对APP开发对应的数据采集SDK工具包（Software Development Kit，软件开发工具包），向用户发布安装有该数据采集SDK工具包的APP。传统方式中，如果要实现多款APP用户行为数据的采集，APP开发商需要手动针对每款APP分别开发对应的数据采集SDK工具包，开发成本和维护成本高，且开发效率低。

发明内容

根据本申请公开的各种实施例，提供一种数据采集工具包定制方法、装置、终端和存储介质。

一种数据采集工具包定制方法，包括:

获取多个数据采集脚本；所述数据采集脚本具有对应的多个数据采集项；

通过数据采集配置页面显示所述数据采集项；

获取对数据采集项的选定操作，根据选定操作获取被选定的数据采集项以及与被选定的数据采集项对应的配置参数；

根据被选定的数据采集项提取对应的数据采集脚本，利用配置参数对提取到的数据采集脚本进行更新，对更新后的数据采集脚本进行编译；及

对编译后的数据采集脚本进行混淆加密，对加密后的数据采集脚本进行封装以得到定制的数据采集工具包。

一种数据采集工具包定制装置，所述装置包括:

页面展示模块，用于获取多个数据采集脚本；所述数据采集脚本具有对应的多个数据采集项；通过数据采集配置页面显示所述数据采集项；

参数配置模块，用于获取对数据采集项的选定操作，根据选定操作获取被选定的数据采集项以及与被选定的数据采集项对应的配置参数；

脚本更新模块，用于根据被选定的数据采集项提取对应的数据采集脚本，利用配置参数对提取到的数据采集脚本进行更新，对更新后的数据采集脚本进行编译；及

混淆加密模块，用于对编译后的数据采集脚本进行混淆加密，对加密后的数据采集脚本进行封装以得到定制的数据采集工具包。

一种终端，包括存储器和处理器，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处理器执行时，使得所述处理器执行以下步骤：

获取多个数据采集脚本；所述数据采集脚本具有对应的多个数据采集项；

通过数据采集配置页面显示所述数据采集项；

获取对数据采集项的选定操作，根据选定操作获取被选定的数据采集项以及与被选定的数据采集项对应的配置参数；

根据被选定的数据采集项提取对应的数据采集脚本，利用配置参数对提取到的数据采集脚本进行更新，对更新后的数据采集脚本进行编译；及

对编译后的数据采集脚本进行混淆加密，对加密后的数据采集脚本进行封装以得到定制的数据采集工具包。

一个或多个存储有计算机可读指令的非易失性可读存储介质，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：

获取多个数据采集脚本；所述数据采集脚本具有对应的多个数据采集项；

通过数据采集配置页面显示所述数据采集项；

获取对数据采集项的选定操作，根据选定操作获取被选定的数据采集项以及与被选定的数据采集项对应的配置参数；

根据被选定的数据采集项提取对应的数据采集脚本，利用配置参数对提取到的数据采集脚本进行更新，对更新后的数据采集脚本进行编译；及

对编译后的数据采集脚本进行混淆加密，对加密后的数据采集脚本进行封装以得到定制的数据采集工具包。

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

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 为一个实施例中数据采集工具包定制方法的应用环境图；

图 2 为一个实施例中数据采集工具包定制方法的流程图；

图 3 为一个实施例中数据采集工具包定制装置的结构图；

图 4 为一个实施例中终端的框图。

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，此处描述的具体实施例仅仅用以解释本申请，并不用于限定本申请。

可以理解，本申请所使用的术语“第一”、“第二”等可在本文中用于描述各种元件，但这些元件不受这些术语限制。这些术语仅用于将第一个元件与另一个元件区分。举例来说，在不脱离本申请的范围的情况下，可以将第一客户端称为第二客户端，且类似地，可将第二客户端称为第一客户端。第一客户端和第二客户端两者都是客户端，但其不是同一客户端。

本申请提供一种数据采集工具包定制方法，可以应用于如图 1 所示的应用环境中。其中，定制终端 102 通过网络与用户终端 104 连接。用户终端 104 可以是智能手机、平板电脑、笔记本电脑、POS 机、车载电脑中的至少一种，但并不局限于此。用户终端 104 上安装有应用程序。定制终端 102 用于生成不同应用程序对应的数据采集工具包，并向对应的用户终端 104 发布安装有该数据采集工具包的应用程序。定制终端 102 预存储了多个数据采集脚本，每个数据采集脚本具有对应的一个或多个数据采集项。定制终端 102 将多个数据采集项以数据采集配置页面的形式显示。基于显示的数据采集配置页面，应用程序开发商可以根据开发需求在定制终端 102 对数据采集项进行选择，并输入与选定的数据采集项对应的配置参数，如自定义功能函数，或者包括多个预设标识的定制数组等。定制终端 102 根据被选定的数据采集项提取对应的数据采集脚本，利用配置参数对提取到的数据采集脚本进行更新。例如，定制终端 102 利用自定义功能函数对提取到的数据采集脚本进行填充；或对定制数组中的预设标识进行循环遍历，将定制数组中的预设标识插入数据采集脚本中的类名，得到多个更新后的数据采集脚本。定制终端 102 对更新后的数据采集脚本进行编译，并对编译后的数据采集脚本进行混淆加密和封装，得到定制的数据采集工具包。当需要针对不同应用程序开发数据采集工具包时，只需变更对数据采集项选择即可，可以提高数据采集工具包开发效率。

在一个实施例中，如图 2 所示，提供了一种数据采集工具包定制方法，以该方法应用于定制终端为例进行说明，具体包括以下步骤：

步骤 202，获取多个数据采集脚本；数据采集脚本具有对应的多个数据采集项。

当需要针对某款应用程序定制对应的数据采集工具包时，应用程序开发商可以通过在定制终端的命令行窗口输入预设的定制命令，向定制终端发起数据采集工具包定制请求。定制终端可以是多款应用程序的控制终端。这里的应用程序可以是金融应用、理财应用、电商应用等需要进行网络交易的敏感应用。根据定制终端上运行的操作系统不同，预先编写的定制命令可以不同。例如，在 andriod（安卓）平台，定制命令可以是 gradle 命

令（一种项目自动化构建工具）。gradle 命令是应用程序开发商利用 groovy 语言（一种编程语言）编写的能够在构建工具 gradle 中运行的脚本。可以理解，还可以采用其他方式触发定制终端对数据采集工具包进行定制，对此不作限制。

定制终端根据数据采集工具包定制请求获取预先封装的基础的数据采集工具包。基础的数据采集工具包可以是应用程序开发商自行开发的。基础的数据采集工具包也可以由第三方开发，并将开发的数据采集工具包以接口的形式对外开放，以供定制终端调用。

数据采集工具包用于对应用程序中的目标数据进行采集。目标数据可以是应用程序页面对应的视图信息或用户行为数据等，如视图名称，视图属性，用户对应用程序页面的触摸操作和访问频率等。根据数据采集脚本能够实现的功能，每个数据采集脚本具有对应的脚本标识。脚本标识包括数据采集项以及对数据采集项的处理操作。也就是说，每个数据采集脚本与一个或多个数据采集项相关联。例如，数据采集脚本 SDK.x 用于对应用程序页面中的视图属性进行采集，则可以用“视图属性采集”标识该数据采集脚本；数据采集脚本 SDK.y 用于采集应用程序页面中视图对应的触摸事件，并对触摸事件进行拦截，则可以用“触摸事件拦截”标识该数据采集脚本。

步骤 204，通过数据采集配置页面显示所述数据采集项。

步骤 206，获取对数据采集项的选定操作，根据选定操作获取被选定的数据采集项以及与被选定的数据采集项对应的配置参数。

定制终端获取多个数据采集脚本对应的脚本标识，并通过数据采集配置页面对多个脚本标识进行展示。应用程序开发商可以根据数据采集配置页面展示的脚本标识对每个数据采集脚本的功能进行了解后，在数据采集配置页面对一个或多个脚本标识进行选择。

预先封装的基础的数据采集工具包中的数据采集脚本功能较为单一，为了满足应用程序开发商对每个数据采集项的个性化采集需求，定制终端在数据采集配置页面对每个脚本标识添加对应的脚本变更输入框。当对某个数据采集项具有特定的数据采集需求时，应用程序开发商可以在该数据采集项对应的脚本变更输入框录入相应的配置参数。容易理解，应用程序开发商也可以在数据采集配置页面以文件上传的方式对配置参数进行录入，对此不作限定。配置参数可以是应用程序开发商补充的功能函数或者自定义的字符串等。

步骤 208，根据被选定的数据采集项提取对应的数据采集脚本，利用配置参数对提取到的数据采集脚本进行更新，对更新后的数据采集脚本进行编译。

定制终端根据应用程序开发商在数据采集配置页面选定的一个或多个脚本标识，在预先封装的基础的数据采集工具包中提取对应的数据采集脚本。若选定的一个或多个脚本标识具有对应的配置参数，则定制终端利用配置参数对相应的数据采集脚本进行更新。

在一个实施例中，配置参数包括自定义功能函数，利用配置参数对提取到的数据采集脚本进行更新的步骤，包括：利用自定义功能函数对提取到的数据采集脚本进行填充。

当预先封装的数据采集工具包中的某个数据采集脚本无法达到应用程序开发商的采集需求时，应用程序开发商可以在数据采集配置页面将自定义功能函数作为配置参数录入

对应的脚本变更输入框。自定义功能函数可以是应用程序开发商根据期望数据采集脚本能够实现的其他功能预先编写的。配置参数还包括自定义功能函数对应的添加方式，添加方式包括插入和替换。当添加方式为插入时，配置参数还包括自定义功能函数的插入位置；当添加方式为替换时，配置参数还包括自定义功能函数的替换对象。

当自定义功能函数的添加方式为插入时，定制终端根据自定义功能函数的插入位置，将自定义功能函数插入对应的数据采集脚本。例如，在上述举例中，应用程序开发商期望数据采集脚本 SDK.y 可以对触摸事件进行上报，则可以将触摸事件上报功能函数插入至数据采集脚本 SDK.y。当自定义功能函数的添加方式为替换时，定制终端根据自定义功能函数的替换对象，在相应的数据采集脚本中将替换对象删除，在该替换对象所在位置插入自定义功能函数。应用程序开发商只需根据开发需求编写新增功能的自定义功能函数，将该自定义功能函数填充至基础的数据采集脚本，即可实现对数据采集脚本的更新，而无需对整个数据采集脚本进行修改，减少应用程序开发商开发量，进而可以提高数据采集工具包开发效率。

在一个实施例中，配置参数包括定制数组，定制数组包括多个预设标识；利用配置参数对提取到的数据采集脚本进行更新的步骤，包括：对定制数组中的预设标识进行循环遍历，将定制数组中的预设标识插入数据采集脚本中的类名，得到多个更新后的数据采集脚本。

传统方式中，在利用数据采集工具包对用户行为数据进行采集之前，应用程序开发商需要预先针对不同的应用程序分别开发对应的数据采集工具包。但事实上，同一应用程序开发商开发的多款应用程序的实现逻辑相同或相似，使得多款应用程序分别对应的数据采集工具包相似。例如，数据采集工具包包括多个数据采集脚本。每个数据采集脚本包括一个或多个对象类，每个对象类具有对应的类名；每个对象类包括一个或多个函数，每个函数具有对应的函数名。部分应用程序对应的数据采集工具包的区别仅在于类名或函数名。若需要进行目标数据采集的应用程序较多，则应用程序开发商需要进行大量的重复动作，浪费人力，也降低了数据采集效率。

为了减少重复操作，定制终端允许应用程序开发商在数据采集页面将定制数组作为配置参数进行录入，利用定制数组批量定制数据采集工具包。数据采集工具包与实现逻辑相同或相似的多款应用程序相对应。具体的，定制数组为一维数组，将一个或多个预设标识作为数组元素进行记录。预设标识可以是字符或字符串等。

定制终端将定制数组中第一个数组元素作为前缀或后缀添加至相应数据采集脚本中全部类名中，以对数据采集脚本进行更新。定制终端对定制数组进行遍历，按照上述方式将下一个数组元素作为前缀或后缀添加至相应数据采集脚本中全部类名中，以此类推，批量获得多个更新后的数据采集脚本。定制终端对更新后的数据采集脚本可以获取多个定制的数据采集工具包。容易理解，获得数据采集工具包的数量与定制数组中数组元素的数量相同。应用程序开发商可以通过变更定制数组的数组元素，对需要生成的定制的数据采集

工具包的数量自由调整，提高数据采集工具包的定制效率。例如，假设定制数组为[ABC，DEF，GHI]，在相应数据采集脚本中所有类名前添加前缀 ABC，编译后获得定制的数据采集工具包 X；在相应数据采集脚本中所有类名前添加前缀 DEF，编译后获得定制的数据采集工具包 Y；在相应数据采集脚本中所有类名前添加前缀 GHI，编译后获得定制的数据采集工具包 Z。

步骤 210，对编译后的数据采集脚本进行混淆加密，对加密后的数据采集脚本进行封装以得到定制的数据采集工具包。

定制终端对更新后的数据采集脚本进行编译，执行封装命令，得到定制的数据采集工具包。定制终端将定制的数据采集工具包安装至对应的应用程序，将安装有定制的数据采集工具包的应用程序发布至用户终端，以使用户终端在相应的应用程序中安装该定制的数据采集工具包。在一个实施例中，由于需要进行目标数据采集的应用程序可能是涉及网络交易的敏感应用，为了减少敏感数据的泄漏风险，在对编译后的数据采集脚本进行封装之前，定制终端对编译后的数据采集脚本混淆加密，以提高数据采集工具包的安全性。

本实施例中，通过数据采集配置页面对多个数据采集项进行展示，使得应用程序开发商可以在数据采集配置页面对数据采集项进行选择并输入与选定的数据采集项对应的配置参数；根据被选定的数据采集项，可以提取对应的数据采集脚本；利用配置参数对提取到的数据采集脚本进行更新，可以满足个性化的数据采集需求；在对更新后的数据采集脚本进行编译，混淆加密和封装后，可以得到安全性高的数据采集工具包。采用页面配置的方式选定数据采集项，使得无需重新编写代码，只需添加或删除易读的数据采集项即可实现数据采集工具包定制，降低了数据采集工具包的开发门槛和开发成本。当需要针对其他某款应用程序开发数据采集工具包时，只需变更对数据采集项的选择即可，可以提高数据采集工具包开发效率。

在一个实施例中，对编译后的数据采集脚本进行混淆加密的步骤，包括：获取数据采集脚本中的敏感字段，计算敏感字段对应的二进制数值；按照预设规则将数据采集脚本中的类名或函数名进行拼接，得到敏感字段对应的盐值，计算盐值对应的二进制数值；对敏感字段对应的二进制数值和盐值对应的二进制数值进行预设逻辑运算，得到混淆字符串；通过将混淆字符串对相应敏感字段进行替换，对数据采集脚本进行混淆加密。

为了提高数据采集工具包中的敏感字段的安全性，应用程序开发商通常采用盐值加密的方式对敏感字段进行加密。敏感字段包括用户名、密码、密钥、域名、IP地址（Internet Protocol，网络协议）、CA（Certification Authority，认证）数字证书和数字签名等。盐值加密是指给待加密的敏感字段关联一个n位随机序列，该n位随机序列即为“盐值”。传统的盐值加密方式中，利用系统方法将敏感字段转化为可读的字符串。但采用系统方法难以保证字符串转换成功率，一旦转换失败，则需要更换盐值，降低加密效率。更重要的是，当字符串转换失败时，敏感字段关联的盐值面临着存储的问题，如果将盐值以明文的方式内置在数据采集脚本中则不利于加密效果，敏感字段的安全性仍不高。

为了解决上述问题，定制终端获取更新后的采集脚本中的敏感字段，动态生成敏感字段对应的盐值，根据敏感字段及对应盐值生成混淆字符串，利用混淆字符串对敏感字段进行加密。具体的，定制终端将敏感字段转化为二进制数值。定制终端获取数据采集脚本中的多个类名和函数名，按照预设规则将获取到的类名和函数名进行拼接，得到随机字符串，对随机字符串进行哈希变换，生成敏感字段对应的盐值。定制终端将敏感字段对应的盐值转化为二进制数值，对敏感字段对应的二进制数值和盐值对应的二进制数值分别进行加密，对加密后的敏感字段对应的二进制数值和盐值对应的二进制数值进行预设逻辑运算，得到目标二进制数值。预设逻辑运算包括异或运算。定制终端将目标二进制数值转化为十六进制数值，将该十六进制数值作为混淆字符串对敏感字段进行替换。将包含敏感字段和盐值的数据信息的混淆字符串内置在数据采集脚本中，可以解决敏感字段对应盐值的存储问题，且混淆字符串的获取只需简单的进制转换，可以解决字符串转换失败的问题，从而可以提高对敏感字段的加密效率。

定制终端将混淆加密后的数据采集工具包发布至用户终端。当合法用户在用户终端使用安装有该数据采集工具包的应用程序时，用户终端获取数据采集脚本中的混淆字符串，对混淆字符串进行解密，将混淆字符串还原为敏感字段。具体的，用户终端将数据采集脚本中的混淆字符串转换为二进制数值。用户终端按照上述方式生成混淆字符串对应的盐值，即获取数据采集脚本中的多个类名和函数名，按照预设规则将获取到的类名和函数名进行拼接，得到随机字符串，对随机字符串进行哈希变换，生成混淆字符串对应的盐值。由于均是按照预设规则利用数据采集脚本中的类名和函数名生成，对混淆字符串解密时生成的盐值与对敏感字段加密时生成的盐值相同。用户终端将混淆字符串对应的盐值转化为二进制数值，对混淆字符串对应的二进制数值和盐值对应的二进制数值进行预设逻辑运算，得到目标二进制数值。预设逻辑运算包括异或运算。用户终端将目标二进制数值转化为十六进制数值。可以理解，该十六进制数值即为敏感字段。敏感字段的获取只需简单的进制转换，可以提高对混淆字符串的解密效率。

本实施例中，混淆加密过程利用数据采集脚本中的类名或函数名动态产生盐值，可以减少盐值被泄漏的风险。将包含敏感字段和盐值的数据信息的混淆字符串内置在数据采集脚本中，可以解决敏感字段对应盐值的存储问题。混淆字符串的获取只需简单的进制转换，可以解决字符串转换失败的问题，进而可以提高对敏感字段的加密效率。

在一个实施例中，对编译后的数据采集脚本进行混淆加密的步骤，包括：通过函数指针将数据采集脚本中的代码封装至预设主函数，预设主函数包括多条输入输出类型转换语句；利用输入输出类型转换语句将数据采集脚本的输入类型转换为第一固定类型，输出类型转换为第二固定类型，以对数据采集脚本进行混淆加密。

传统方式中，利用混淆加密工具对数据采集脚本进行混淆加密。但现有混淆加密工具对数据采集脚本的混淆加密强度不够，很容易被攻击者破解，造成客户信息等敏感数据的泄漏。例如，使用 escape 混淆加密后的数据采集脚本，只需要使用 unescape 函数即可还

原出源数据采集脚本，现有的混淆加密工具难以保证数据采集脚本的安全。

为了解决上述问题，定制终端对更新后的数据采集脚本的输入类型和输出类型进行混淆。具体的，定制终端对更新后的数据采集脚本进行编译后，将数据采集脚本中的代码封装至预设主函数中。预设主函数可以是基于预设编程语言实现的主函数。容易理解，根据操作系统的不同，数据采集脚本中的代码采用的编程语言可以不同。例如，在 andriod 平台，可以采用 Java 语言（一种编程语言）编写数据采集脚本；在 IOS（iPhone Operating System，苹果操作系统）平台，可以采用 Objective-C 语言（一种编程语言）编写数据采集脚本；在混合应用平台，可以采用 JS 语言（JavaScript，直译式脚本语言）编写数据采集脚本。预设编程语言可以是区别于开发数据采集脚本时采用的编程语言，如 C 语言或者 C++语言（一种编程语言）等。

预设主函数包括多条输入输出类型转换语句。输入输出类型转换语句用于将数据采集脚本中多样化的数据输入类型和输出类型分别转换为第一固定类型和第二固定类型。输入输出类型转换语句可以通过在预设主函数定义结构体来实现。具体的，预设主函数包括多个结构体，每个结构体中声明了多个函数指针，将函数指针指向数据采集脚本中的一个或多个对象类。每个结构体具有对应的结构体变量，结构体变量数据类型为第一固定类型或第二固定类型。第一固定类型和第二固定类型分别可以是 int 型（整数型），long 型（长整数型）或者 String（字符串类型）等。当定制终端对更新后的数据采集脚本进行编译时，预设主函数利用函数指针调用对应数据采集脚本中的代码，从而将数据采集脚本中的代码封装至预设主函数中。定制终端通过执行预设主函数将数据采集脚本的输入类型和输出类型转换为固定的类型，从而对数据采集脚本的输入类型和输出类型进行混淆，且该混淆为不可逆混淆。可以理解，输入输出类型转换语句也可以通过其他方式实现，对此不作限制。

本实施例中，对数据采集脚本中每个对象类的输入类型和输出类型进行混淆，且该混淆为不可逆混淆，攻击者通过反编译得到的输入类型和输出类型均为数字或字母，攻击者无从知晓数据采集脚本的实际输入类型和输出类型，从而也就无法获知数据采集脚本中各个函数的实际功能，增加了数据采集脚本的破解成本。

在一个实施例中，对编译后的数据采集脚本进行混淆加密的步骤，包括：获取数据采集脚本中的多个具有功能含义的函数名；获取函数名对应的自定义字符串；通过将自定义字符串对相应的函数名进行替换，对函数名的功能含义进行隐藏，以对数据采集脚本进行混淆加密。

为了提高数据采集工具包的安全性，定制终端将数据采集脚本中具有明确的功能含义的函数名替换为无实际意义的自定义字符串。自定义字符串可以是应用程序开发商在数据采集配置页面录入的配置参数，也可以是定制终端按照预设规则随机生成的。自定义字符串可以是由安全系统较高的预设字符和/或汉字组合而成，如“#”，“@”，“*”等。替换方式可以是在数据采集脚本中添加宏定义语句等，宏定义语句包括被替换函数名和对应的自定义字符串，如#define SafeHideInvoke（x） a#，从而将 SafeHideInvoke（x）函数名替换

为自定义字符串 a#。对函数名进行替换后，攻击者通过反编译得到的函数名均为无实际意义的字符或汉字，难以根据函数名猜测函数功能，进而可以增加数据采集脚本的破解成本。

需要说明的是，数据采集脚本中具有明确的功能含义的函数名对应的字符长度普遍较长。本方案中，自定义字符串的长度小于对应函数名的长度，从而采用自定义字符串替换数据采集脚本中的函数名还可以对数据采集脚本起到代码压缩的作用，减少对定制终端资源的占用。

在一个实施例中，对编译后的数据采集脚本进行混淆加密的步骤，包括：查询数据采集脚本中函数调用语句；函数调用语句包括调用函数的函数名及被调用函数的函数名；根据被调用函数的函数名查询被调用函数的函数体，通过在被调用函数的函数体添加预设关键字，对调用函数和被调用函数之间的调用关系进行隐藏，以对数据采集脚本进行混淆加密。

为了提高数据采集工具包的安全性，定制终端对数据采集脚本中多个函数之间的调用关系进行隐藏。具体的，定制终端在更新后的数据采集脚本中查询是否存在函数调用语句。函数调用语句包括调用函数的函数名和被调用函数的函数名，如函数 static sum(){ c=fun(a,b);}中，sum()为调用函数的函数名；fun(a,b)为被调用函数的函数名。定制终端通过在被调用函数对应函数体的返回类型项前添加预设关键字，对该调用函数和被调用函数之间的调用关系进行隐藏。

预设关键字可以是 inline 关键字（一种关键字）。通过在被调用函数对应函数体的返回类型项前添加 inline 关键字，可以将被调用函数指定为对应调用函数的内联函数。当执行调用函数时，内联函数可以直接插入调用函数，从而对被调用函数进行执行，区别于传统的采用直接调用的方式对被调用函数进行执行，从而可以对调用函数和被调用函数之间的调用关系进行隐藏。

本实施例中，对函数之间的调用关系进行隐藏后，攻击者通过反编译得到的多个孤立的函数，无从知晓多个函数之间的逻辑关系，从而难以确定每个函数对数据采集脚本的功能作用，进而可以增加数据采集脚本的破解成本。

应当理解的是，虽然图 2 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，图 2 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。

在一个实施例中，如图 3 所示，提供了一种数据采集工具包定制装置，包括：页面展示模块 302，参数配置模块 304，脚本更新模块 306 和混淆加密模块 308，其中：

页面展示模块 302，用于获取多个数据采集脚本；数据采集脚本具有对应的多个数据

采集项；通过数据采集配置页面显示数据采集项。

参数配置模块 304，用于获取对数据采集项的选定操作，根据选定操作获取被选定的数据采集项以及与被选定的数据采集项对应的配置参数。

脚本更新模块 306，用于根据被选定的数据采集项，从数据采集工具包中提取对应的数据采集脚本，利用配置参数对提取到的数据采集脚本进行更新，对更新后的数据采集脚本进行编译。

混淆加密模块 308，用于对编译后的数据采集脚本进行混淆加密，对加密后的数据采集脚本进行封装以得到定制的数据采集工具包。

在一个实施例中，混淆加密模块 308 还用于获取数据采集脚本中的敏感字段，计算敏感字段对应的二进制数值；按照预设规则将数据采集脚本中的类名或函数名进行拼接，得到敏感字段对应的盐值，计算盐值对应的二进制数值；对敏感字段对应的二进制数值和盐值对应的二进制数值进行预设逻辑运算，得到混淆字符串；通过将混淆字符串对相应敏感字段进行替换，对数据采集脚本进行混淆加密。

上述的数据采集工具包定制装置可以实现为一种计算机可读指令的形式，计算机可读指令可在如图 4 所示的终端上运行。

在一个实施例中，提供了一种终端的内部框图，如图 4 所示，该终端包括通过系统总线连接的处理器、内存储器、非易失性存储介质、网络接口、显示屏和输入装置。其中，该处理器用于提供计算和控制能力，支撑整个终端的运行。该定制终端的非易失性存储介质存储有操作系统和计算机可读指令，该计算机可读指令可被处理器所执行以实现以下各实施例所提供的一种数据采集工具包定制方法。终端中的内存储器为非易失性存储介质中的操作系统和计算机可读指令的运行提供环境。网络接口用于连接到网络进行通信。显示屏用于显示各种界面，比如，可显示数据采集配置页面。输入装置可以是显示屏上覆盖的触摸层，也可以是电子设备外壳上设置的按键、轨迹球或触控板，也可以是外接的键盘、触控板或鼠标等，用于供输入配置参数和各种指令。处理器执行计算机可读指令时可以执行以下步骤：获取多个数据采集脚本；数据采集脚本具有对应的多个数据采集项；通过数据采集配置页面显示数据采集项；获取对数据采集项的选定操作，根据选定操作获取被选定的数据采集项以及与被选定的数据采集项对应的配置参数；根据被选定的数据采集项提取对应的数据采集脚本，利用配置参数对提取到的数据采集脚本进行更新，对更新后的数据采集脚本进行编译；及对编译后的数据采集脚本进行混淆加密，对加密后的数据采集脚本进行封装以得到定制的数据采集工具包。

本领域技术人员可以理解，图 4 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终端的限定，具体的终端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。

在一个实施例中，配置参数包括自定义功能函数或定制数组，定制数组包括多个预设

标识；处理器执行计算机可读指令时还执行以下步骤：利用自定义功能函数对提取到的数据采集脚本进行填充；或定制数组中的预设标识进行循环遍历，将定制数组中的预设标识插入数据采集脚本中的类名，得到多个更新后的数据采集脚本。

在一个实施例中，处理器执行计算机可读指令时还执行以下步骤：获取数据采集脚本中的敏感字段，计算敏感字段对应的二进制数值；按照预设规则将数据采集脚本中的类名或函数名进行拼接，得到敏感字段对应的盐值，计算盐值对应的二进制数值；对敏感字段对应的二进制数值和盐值对应的二进制数值进行预设逻辑运算，得到混淆字符串；及通过将混淆字符串对相应敏感字段进行替换，对数据采集脚本进行混淆加密。

在一个实施例中，处理器执行计算机可读指令时还执行以下步骤：通过函数指针将数据采集脚本中的代码封装至预设主函数，预设主函数包括多条输入输出类型转换语句；及利用输入输出类型转换语句将数据采集脚本的输入类型转换为第一固定类型，输出类型转换为第二固定类型，以对数据采集脚本进行混淆加密。

在一个实施例中，处理器执行计算机可读指令时还执行以下步骤：获取数据采集脚本中的多个具有功能含义的函数名；获取函数名对应的自定义字符串；及通过将自定义字符串对相应的函数名进行替换，对函数名的功能含义进行隐藏，以对数据采集脚本进行混淆加密。

在一个实施例中，处理器执行计算机可读指令时还执行以下步骤：查询数据采集脚本中函数调用语句；函数调用语句包括调用函数的函数名及被调用函数的函数名；及根据被调用函数的函数名查询被调用函数的函数体，通过在被调用函数的函数体添加预设关键字，对调用函数和被调用函数之间的调用关系进行隐藏，以对数据采集脚本进行混淆加密。

在一个实施例中，提供了一种一个或多个存储有计算机可读指令的非易失性可读存储介质，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：获取多个数据采集脚本；数据采集脚本具有对应的多个数据采集项；通过数据采集配置页面显示数据采集项；获取对数据采集项的选定操作，根据选定操作获取被选定的数据采集项以及与被选定的数据采集项对应的配置参数；根据被选定的数据采集项提取对应的数据采集脚本，利用配置参数对提取到的数据采集脚本进行更新，对更新后的数据采集脚本进行编译；及对编译后的数据采集脚本进行混淆加密，对加密后的数据采集脚本进行封装以得到定制的数据采集工具包。

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，前述计算机可读指令可存储于一非易失性计算机可读取存储介质中，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，前述存储介质可为磁碟、光盘、只读存储记忆体（Read-Only Memory，ROM）等非易失性可读存储介质。

以上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、一种数据采集工具包定制方法，包括:

获取多个数据采集脚本；所述数据采集脚本具有对应的多个数据采集项；

通过数据采集配置页面显示所述数据采集项；

获取对数据采集项的选定操作，根据选定操作获取被选定的数据采集项以及与被选定的数据采集项对应的配置参数；

根据被选定的数据采集项提取对应的数据采集脚本，利用配置参数对提取到的数据采集脚本进行更新，对更新后的数据采集脚本进行编译；及

对编译后的数据采集脚本进行混淆加密，对加密后的数据采集脚本进行封装以得到定制的数据采集工具包。

2、根据权利要求 1 所述的方法，其特征在于，所述配置参数包括自定义功能函数或定制数组，所述定制数组包括多个预设标识；所述利用配置参数对提取到的数据采集脚本进行更新包括:

利用自定义功能函数对提取到的数据采集脚本进行填充；或

对定制数组中的预设标识进行循环遍历，将定制数组中的预设标识插入数据采集脚本中的类名，得到多个更新后的数据采集脚本。

3、根据权利要求 1 所述的方法，其特征在于，所述对编译后的数据采集脚本进行混淆加密包括:

获取数据采集脚本中的敏感字段，计算所述敏感字段对应的二进制数值；

按照预设规则将数据采集脚本中的类名或函数名进行拼接，得到敏感字段对应的盐值，计算所述盐值对应的二进制数值；

对所述敏感字段对应的二进制数值和所述盐值对应的二进制数值进行预设逻辑运算，得到混淆字符串；及

通过将所述混淆字符串对相应所述敏感字段进行替换，对所述数据采集脚本进行混淆加密。

4、根据权利要求 1 所述的方法，其特征在于，所述对编译后的数据采集脚本进行混淆加密包括:

通过函数指针将数据采集脚本中的代码封装至预设主函数，所述预设主函数包括多条输入输出类型转换语句；及

利用所述输入输出类型转换语句将数据采集脚本的输入类型转换为第一固定类型，输出类型转换为第二固定类型，以对所述数据采集脚本进行混淆加密。

5、根据权利要求 1 所述的方法，其特征在于，所述对编译后的数据采集脚本进行混淆加密包括:

获取数据采集脚本中的多个具有功能含义的函数名；及

获取函数名对应的自定义字符串；通过将自定义字符串对相应的函数名进行替换，对

所述函数名的功能含义进行隐藏，以对所述数据采集脚本进行混淆加密。

6、根据权利要求 1 所述的方法，其特征在于，所述对编译后的数据采集脚本进行混淆加密包括:

查询数据采集脚本中函数调用语句；所述函数调用语句包括调用函数的函数名及被调用函数的函数名；及

根据被调用函数的函数名查询被调用函数的函数体，通过在所述被调用函数的函数体添加预设关键字，对调用函数和被调用函数之间的调用关系进行隐藏，以对所述数据采集脚本进行混淆加密。

7、一种数据采集工具包定制装置，包括:

页面展示模块，用于获取多个数据采集脚本；所述数据采集脚本具有对应的多个数据采集项；通过数据采集配置页面显示所述数据采集项；

参数配置模块，用于获取对数据采集项的选定操作，根据选定操作获取被选定的数据采集项以及与被选定的数据采集项对应的配置参数；

脚本更新模块，用于根据被选定的数据采集项提取对应的数据采集脚本，利用配置参数对提取到的数据采集脚本进行更新，对更新后的数据采集脚本进行编译；及

混淆加密模块，用于对编译后的数据采集脚本进行混淆加密，对加密后的数据采集脚本进行封装以得到定制的数据采集工具包。

8、根据权利要求 7 所述的装置，其特征在于，所述混淆加密模块还用于获取数据采集脚本中的敏感字段，计算所述敏感字段对应的二进制数值；按照预设规则将数据采集脚本中的类名或函数名进行拼接，得到敏感字段对应的盐值，计算所述盐值对应的二进制数值；对所述敏感字段对应的二进制数值和所述盐值对应的二进制数值进行预设逻辑运算，得到混淆字符串；及通过将所述混淆字符串对相应所述敏感字段进行替换，对所述数据采集脚本进行混淆加密。

9、一种终端，包括存储器及处理器，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处理器执行时，使得所述处理器执行以下步骤:

获取多个数据采集脚本；所述数据采集脚本具有对应的多个数据采集项；

通过数据采集配置页面显示所述数据采集项；

获取对数据采集项的选定操作，根据选定操作获取被选定的数据采集项以及与被选定的数据采集项对应的配置参数；

根据被选定的数据采集项提取对应的数据采集脚本，利用配置参数对提取到的数据采集脚本进行更新，对更新后的数据采集脚本进行编译；及

对编译后的数据采集脚本进行混淆加密，对加密后的数据采集脚本进行封装以得到定制的数据采集工具包。

10、根据权利要求 9 所述的终端，其特征在于，所述配置参数包括自定义功能函数或定制数组，所述定制数组包括多个预设标识；所述处理器执行所述计算机可读指令时还

执行以下步骤:

利用自定义功能函数对提取到的数据采集脚本进行填充;或

对定制数组中的预设标识进行循环遍历,将定制数组中的预设标识插入数据采集脚本中的类名,得到多个更新后的数据采集脚本。

11、根据权利要求9所述的终端,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:

获取数据采集脚本中的敏感字段,计算所述敏感字段对应的二进制数值;

按照预设规则将数据采集脚本中的类名或函数名进行拼接,得到敏感字段对应的盐值,计算所述盐值对应的二进制数值;

对所述敏感字段对应的二进制数值和所述盐值对应的二进制数值进行预设逻辑运算,得到混淆字符串;及

通过将所述混淆字符串对相应所述敏感字段进行替换,对所述数据采集脚本进行混淆加密。

12、根据权利要求9所述的终端,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:

通过函数指针将数据采集脚本中的代码封装至预设主函数,所述预设主函数包括多条输入输出类型转换语句;及

利用所述输入输出类型转换语句将数据采集脚本的输入类型转换为第一固定类型,输出类型转换为第二固定类型,以对所述数据采集脚本进行混淆加密。

13、根据权利要求9所述的终端,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:

获取数据采集脚本中的多个具有功能含义的函数名;及

获取函数名对应的自定义字符串;通过将自定义字符串对相应的函数名进行替换,对所述函数名的功能含义进行隐藏,以对所述数据采集脚本进行混淆加密。

14、根据权利要求9所述的终端,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:

查询数据采集脚本中函数调用语句;所述函数调用语句包括调用函数的函数名及被调用函数的函数名;及

根据被调用函数的函数名查询被调用函数的函数体,通过在所述被调用函数的函数体添加预设关键字,对调用函数和被调用函数之间的调用关系进行隐藏,以对所述数据采集脚本进行混淆加密。

15、一个或多个存储有计算机可读指令的非易失性可读存储介质,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:

获取多个数据采集脚本;所述数据采集脚本具有对应的多个数据采集项;

通过数据采集配置页面显示所述数据采集项;

获取对数据采集项的选定操作，根据选定操作获取被选定的数据采集项以及与被选定的数据采集项对应的配置参数；

根据被选定的数据采集项提取对应的数据采集脚本，利用配置参数对提取到的数据采集脚本进行更新，对更新后的数据采集脚本进行编译；及

对编译后的数据采集脚本进行混淆加密，对加密后的数据采集脚本进行封装以得到定制的数据采集工具包。

16、根据权利要求 15 所述的存储介质，其特征在于，所述配置参数包括自定义功能函数或定制数组，所述定制数组包括多个预设标识；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：

利用自定义功能函数对提取到的数据采集脚本进行填充；或

对定制数组中的预设标识进行循环遍历，将定制数组中的预设标识插入数据采集脚本中的类名，得到多个更新后的数据采集脚本。

17、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还执行以下步骤：

获取数据采集脚本中的敏感字段，计算所述敏感字段对应的二进制数值；

按照预设规则将数据采集脚本中的类名或函数名进行拼接，得到敏感字段对应的盐值，计算所述盐值对应的二进制数值；

对所述敏感字段对应的二进制数值和所述盐值对应的二进制数值进行预设逻辑运算，得到混淆字符串；及

通过将所述混淆字符串对相应所述敏感字段进行替换，对所述数据采集脚本进行混淆加密。

18、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还执行以下步骤：

通过函数指针将数据采集脚本中的代码封装至预设主函数，所述预设主函数包括多条输入输出类型转换语句；及

利用所述输入输出类型转换语句将数据采集脚本的输入类型转换为第一固定类型，输出类型转换为第二固定类型，以对所述数据采集脚本进行混淆加密。

19、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还执行以下步骤：

获取数据采集脚本中的多个具有功能含义的函数名；及

获取函数名对应的自定义字符串；通过将自定义字符串对相应的函数名进行替换，对所述函数名的功能含义进行隐藏，以对所述数据采集脚本进行混淆加密。

20、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还执行以下步骤：

查询数据采集脚本中函数调用语句；所述函数调用语句包括调用函数的函数名及被调

用函数的函数名；及

根据被调用函数的函数名查询被调用函数的函数体，通过在所述被调用函数的函数体添加预设关键字，对调用函数和被调用函数之间的调用关系进行隐藏，以对所述数据采集脚本进行混淆加密。

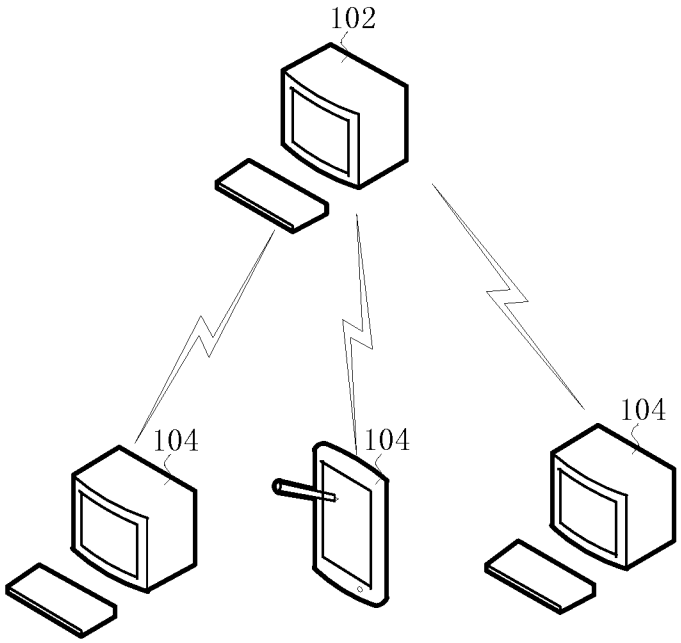

102
104
104
104

图 1

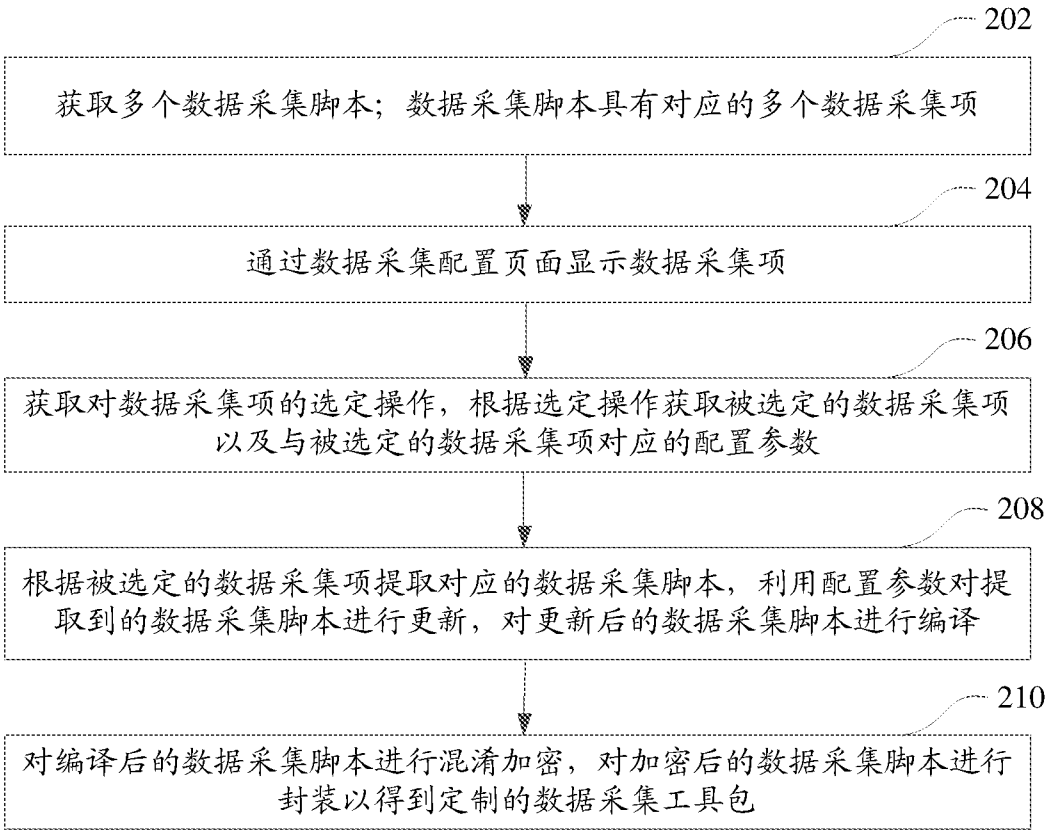

202
获取多个数据采集脚本；数据采集脚本具有对应的多个数据采集项
204
通过数据采集配置页面显示数据采集项
206
获取对数据采集项的选定操作，根据选定操作获取被选定的数据采集项以及与被选定的数据采集项对应的配置参数
208
根据被选定的数据采集项提取对应的数据采集脚本，利用配置参数对提取到的数据采集脚本进行更新，对更新后的数据采集脚本进行编译
210
对编译后的数据采集脚本进行混淆加密，对加密后的数据采集脚本进行封装以得到定制的数据采集工具包

图 2

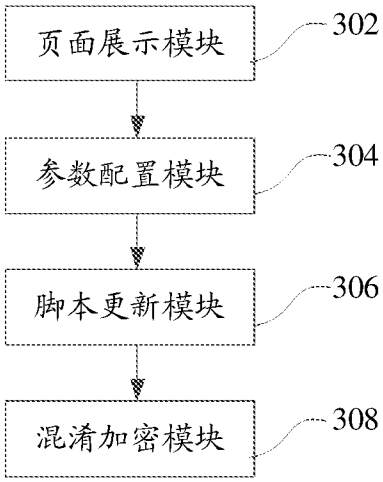

页面展示模块
302
参数配置模块
304
脚本更新模块
306
混淆加密模块
308

图 3

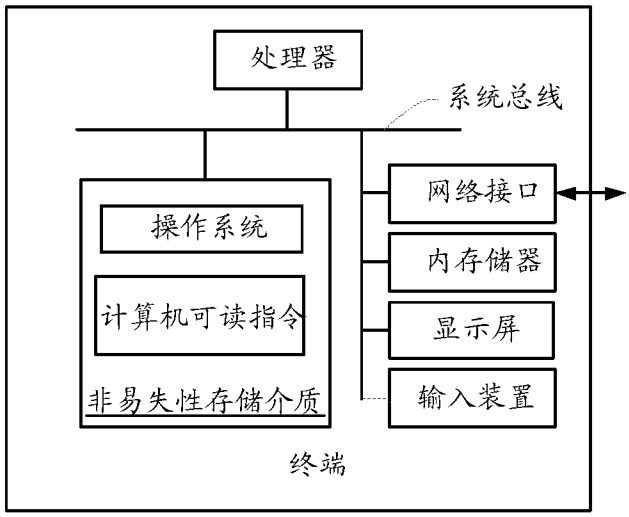

处理器
系统总线
网络接口
操作系统
内存储器
计算机可读指令
显示屏
非易失性存储介质
输入装置
终端

图 4

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2018/076961

A. CLASSIFICATION OF SUBJECT MATTER

G06F 8/30 (2018.01) 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)

SIPOABS; CPRSABS; CNABS; DWPI; CNKI; IEEE: 数据, 采集, 项, 工具包, 定制, 脚本, 混淆, data, collection, kit, customization, script, confusion

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 101267333 A (NANJING LINKAGE TECHNOLOGY CO., LTD.) 17 September 2008 (17.09.2008), claims 1-3	1-20
A	CN 104360837 A (THE THIRD RESEARCH INSTITUTE OF MINISTRY OF PUBLIC SECURITY) 18 February 2015 (18.02.2015), entire document	1-20
A	CN 101833498 A (BEIJING ANTARES TESTING SCIENCE & TECHNOLOGY CO., LTD. et al.) 15 September 2010 (15.09.2010), 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
02 August 2018	09 August 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
HAN, Xianping
Telephone No. (86-10) 62411841

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2018/076961

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 101267333 A	17 September 2008	CN 101267333 B	07 September 2011
CN 104360837 A	18 February 2015	CN 104360837 B	13 October 2017
CN 101833498 A	15 September 2010	CN 101833498 B	06 June 2012

国际检索报告

国际申请号
PCT/CN2018/076961

A. 主题的分类

G06F 8/30(2018.01)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））

SIPOABS;CPRSABS;CNABS;DWPI;CNKI;IEEE:数据，采集，项，工具包，定制，脚本，混淆，data，collection，kit，customization，script，confusion

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 101267333 A（南京联创科技股份有限公司）2008年 9月 17日（2008 - 09 - 17） 权利要求1-3	1-20
A	CN 104360837 A（公安部第三研究所）2015年 2月 18日（2015 - 02 - 18） 全文	1-20
A	CN 101833498 A（北京伟晨豪嘉科技有限责任公司等）2010年 9月 15日（2010 - 09 - 15） 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
2018年 8月 2日	2018年 8月 9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 韩鲜萍 电话号码 (86-10)62411841

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2018/076961

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	101267333	A	2008年 9月 17日	CN	101267333	B	2011年 9月 7日
CN	104360837	A	2015年 2月 18日	CN	104360837	B	2017年 10月 13日
CN	101833498	A	2010年 9月 15日	CN	101833498	B	2012年 6月 6日