(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
2025 年 6 月 19 日 (19.06.2025)

WIPO | PCT

(10) 国际公布号
WO 2025/123744 A1

(51) 国际专利分类号:
G06F 21/62 (2013.01)

(21) 国际申请号: PCT/CN2024/112983

(22) 国际申请日: 2024 年 8 月 19 日 (19.08.2024)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202311709334.1 2023年12月13日 (13.12.2023) CN

(71) 申请人: 中国第一汽车股份有限公司(**CHINA FAW CO., LTD.**) [CN/CN]; 中国吉林省长春市汽车经济技术开发区新红旗大街1号 130011 (CN)。一汽（北京）软件科技有限公司(**FAW BEIJING SOFTWARE TECHNOLOGY CO., LTD.**) [CN/CN]; 中国北京市石景山区石景山路54号院4号楼14层 100043 (CN)。

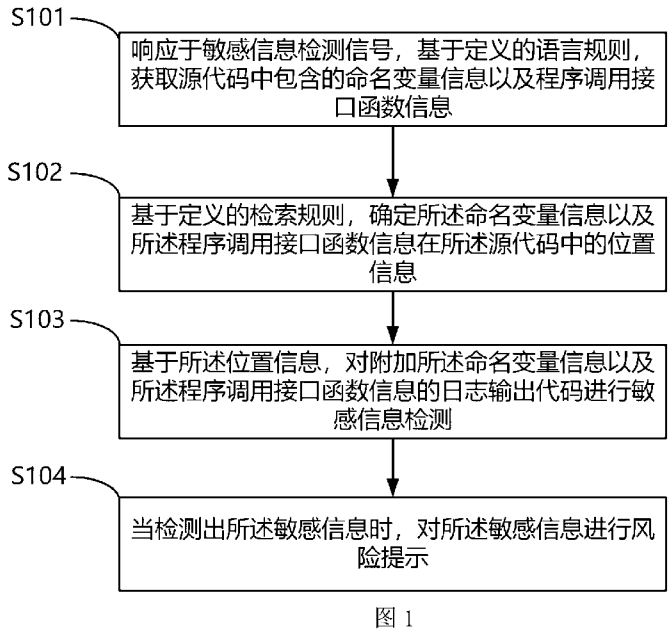
(72) 发明人:尹旻(**YIN, Min**); 中国吉林省长春市汽车经济技术开发区新红旗大街1号 130011 (CN)。王从涛(**WANG, Congtao**); 中国吉林省长春市汽车经济技术开发区新红旗大街1号 130011 (CN)。

(74) 代理人:北京翔宇专利代理有限公司(**BEIJING XIANGYU PATENT AGENCY CO., LTD.**); 中国北京市西城区新兴东巷15号金泰鑫侨大厦6层601室 100044 (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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ,

(54) **Title:** LOG SENSITIVE INFORMATION DETECTION METHOD AND SYSTEM, ELECTRONIC DEVICE, AND STORAGE MEDIUM

(54) 发明名称: 一种日志敏感信息检测方法、系统、电子设备及存储介质



S101 In response to a sensitive information detection signal, on the basis of a defined language rule, acquire naming variable information and program call interface function information comprised in source code

S102 On the basis of a defined retrieval rule, determine position information of the naming variable information and the program call interface function information in the source code

S103 On the basis of the position information, perform sensitive information detection on log output code added with the naming variable information and the program call interface function information

S104 When sensitive information is detected, give a risk alert for the sensitive information

图 1

(57) **Abstract:** Disclosed in the present invention are a log sensitive information detection method and system, an electronic device, and a storage medium. The method comprises: in response to a sensitive information detection signal, on the basis of a defined language rule, acquiring naming variable information and program call interface function information comprised in source code; on the basis of a defined retrieval rule, determining position information of the naming variable information and the program call interface function information in the source code; on the basis of the position information, performing sensitive information detection on log output code added with the naming variable information and the program call interface function information; and when sensitive information is detected, give a risk alert for the sensitive information. The method is independent of the log output form, the position related to sensitive

WO 2025/123744 A1

[见续页]

LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, 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, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, 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))。

information can be detected, and whether there is a potential risk of sensitive information leakage can also be detected; the position of an issue in source code can be identified while the risk is detected, which is conducive to confirming and fixing the issue; the method is independent of dynamic operation of a program, the detection result is more accurate and comprehensive, and the efficiency is higher.

(57) 摘要: 本发明公开了一种日志敏感信息检测方法、系统、电子设备及存储介质，包括响应于敏感信息检测信号，基于定义的语言规则，获取源代码中包含的命名变量信息以及程序调用接口函数信息；基于定义的检索规则，确定命名变量信息以及程序调用接口函数信息在源代码中的位置信息；基于位置信息，对附加命名变量信息以及程序调用接口函数信息的日志输出代码进行敏感信息检测；当检测出敏感信息时，对敏感信息进行风险提示。此方法不依赖日志的输出形式，可以检测出涉及敏感信息位置，并进一步检测是否存续敏感信息泄漏的风险，检测到风险的同时，可明确问题所在源代码位置，便于确认和修复问题，不依赖于程序的动态运行，检测结果更准确、全面且效率更高。

一种日志敏感信息检测方法、系统、电子设备及存储介质

技术领域

本发明涉及信息技术领域，特别是涉及一种日志敏感信息检测方法、系统、电子设备及存储介质。

背景技术

系统日志可用于软件的开发调试以及系统运行分析等，开发过程中开发人员可能会把各种所需的调试信息都打印在日志中，甚至包括一些敏感信息，例如：某些系统的用户名和密码、设备唯一编码、公司内部的用户信息字段等。开发过程中这些数据都是测试数据或者伪造的临时数据，打印在日志中不会造成危害。在完成开发后的正式发布版本中这些日志代码都应该删除或关闭，但由于主观疏忽或流程不完善等因素影响，这些代码没有被正确处理，导致敏感信息可能出现在最终正式版本的系统中，造成信息泄露，带来严重的危害。

针对这种问题，需要对日志文件进行专门的处理，去掉敏感数据，避免泄露，即所谓的脱敏。日志脱敏是信息安全的一个重要方向。通常大型系统的日志脱敏会得到更多的关注，例如金融交易系统、线上购物平台、银行系统等，这些系统的日志相对比较规范、标准，结构化程度高，在系统的子模块之间有清晰的接口和边界，便于检查处理。通过预定规则和模式等检测，可识别出部分可能的信息泄露风险。

CN116383885A 的专利文件公开了一种日志脱敏方法包括：获取目标交易的交易日志，并确定交易日志的日志格式；获取预设字段信息，其中，预设字段信息包括对交易日志进行脱敏的多个字段的信息，多个字段的信息包括：用户信息：基于交易日志的日志格式和预设字段信息，通过 AC 自动机，确定交易日志中待脱敏的字段的位置；基于交易日志中待脱敏的字段的位置，对交易日志进行脱敏，得到脱敏后的交易日志。本发明解决了相关技术中通过正则匹配的方式查找报文日志中的敏感字段，以对报文日志进行同步脱敏，应用系统资源占用率高的技术问题。

上述方案都是从日志分析出发，事后检查信息泄露的风险，并进行补充处理，仍存在较大的信息泄露风险。

因此，本申请提供一种日志敏感信息检测方法以解决上述技术问题。

发明内容

本发明提供一种日志敏感信息检测方法、系统、电子设备及存储介质，能够解决上述提到的至少一个技术问题。

为实现本发明目的提供的一种日志敏感信息检测方法，包括：

响应于敏感信息检测信号，基于定义的语言规则，获取源代码中包含的命名变量信息以及程序调用接口函数信息；

基于定义的检索规则，确定所述命名变量信息以及所述程序调用接口函数信息在所述源代码中的位置信息；

基于所述位置信息，对附加所述命名变量信息以及所述程序调用接口函数信息的日志输出代码进行敏感信息检测；

当检测出所述敏感信息时，对所述敏感信息进行风险提示。

在其中一些具体实施例中，基于所述位置信息，对附加所述命名变量信息以及所述程序调用接口函数信息的日志输出代码进行敏感信息检测后，所述方法还包括：

基于所述日志输出代码，对获取的所述命名变量信息以及所述程序调用接口函数信息进行验证，确认所述日志中是否检测到所述敏感信息。

在其中一些具体实施例中，响应于敏感信息检测信号，基于定义的语言规则，获取源代码中包含的命名变量信息以及程序调用接口函数信息，具体包括：

所述定义的语言规则包括，代码编程规则；

基于所述编程规则，获取源代码中包含的命名变量信息以及程序调用接口函数信息；

其中，所述命名变量信息包括，数据库服务器地址、数据名称、用户名以及用户密码；

所述调用接口函数信息包括，连接数据操作、登录系统以及打开服务。

在其中一些具体实施例中，基于定义的检索规则，确定所述命名变量信息以及所述程序调用接口函数信息在所述源代码中的位置信息，具体包括：

所述定义的检索规则包括：

预置关键字列表，对所述命名变量信息进行检索；

对所述源代码进行扫描，对所述程序调用接口函数信息进行检索；

基于检索到的所述命名变量信息以及所述程序调用接口函数信息，确定二者在所述源代码中的所述位置信息。

在其中一些具体实施例中，基于所述位置信息，对附加所述命名变量信息以及所述程序调用接口函数信息的日志输出代码进行敏感信息检测，具体包括：

基于确定的所述命名变量信息以及所述程序调用接口函数信息在所述源代码中的位置信息，确定与之相关联的所述日志；

对所述日志的输出代码进行敏感信息检测，确认所述日志输出代码是否包含所述敏感信息。

在其中一些具体实施例中，当检测出所述敏感信息时，对所述敏感信息进行风险提示，具体包括：

当检测到所述敏感信息时，对所述敏感信息进行标记；

基于所述标记，进行相应的风险提示。

在其中一些具体实施例中，对所述敏感信息进行风险提示后，所述方法还包括：

对标记的所述敏感信息进行脱敏处理。

基于同一构思，本发明还提供一种日志敏感信息检测系统，包括：

信息获取模块，配置为响应于敏感信息检测信号，基于定义的语言规则，获取源代码中包含的命名变量信息以及程序调用接口函数信息；

位置信息确认模块，配置为基于定义的检索规则，确定所述命名变量信息以及所述程序调用接口函数信息在所述源代码中的位置信息；

敏感信息检测模块，配置为基于所述位置信息，对附加所述命名变量信息以及所述程序调用接口函数信息的日志输出代码进行敏感信息检测；

敏感信息提示模块，配置为当检测出所述敏感信息时，对所述敏感信息进行风险提示。

基于同一构思，本发明还提供一种电子设备，包括：处理器、通信接口、存储器和通信总线，其中，处理器，通信接口，存储器通过通信总线完成相互间的通信；所述存储器中存储有计算机程序，当所述计算机程序被所述处理器执行时，使得所述处理器执行上述日志敏感信息检测方法的步骤。

基于同一构思，本发明还提供一种计算机可读存储介质，其特征在于，其存储有可由电子设备执行的计算机程序，当所述计算机程序在所述电子设备上运行时，使得所述电子设备执行上述日志敏感信息检测方法的步骤。

与现有技术相比，本发明具有以下有益效果：

本发明公开了一种日志敏感信息检测方法、系统、电子设备及存储介质，基于源代码自动检测敏感信息，不依赖日志的输出形式，可以检测出涉及敏感信息位置，并进一步检测是否存续敏感信息泄漏的风险，检测到风险的同时，可明确问题所在源代码位置，便于确认和修复问题，不依赖于程序的动态运行，检测结果更准确、全面且效率更高。

附图说明

图 1 是本发明一种日志敏感信息检测方法在一些具体实施例的流程示意图；

图 2 是本发明一种日志敏感信息检测方法在一些应用中的流程示意图；

图 3 是本发明一种日志敏感信息检测系统在一些具体实施例的结构示意图；

图 4 是本发明一种电子设备在一些具体实施例的结构示意图。

具体实施方式

为了使本申请的目的、技术方案和优点更加清楚，下面将结合附图对本申请作进一步地详细描述，显然，所描述的实施例仅仅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它实施例，都属于本申请保护的范围。

在本申请实施例中使用的术语是仅仅出于描述特定实施例的目的，而非旨在限制本申请。在本申请实施例和所附权利要求书中所使用的单数形式的“一种”、“所述”和“该”也旨在包括多数形式，除非上下文清楚地表示其他含义，“多种”一般包含至少两种。

应当理解，本文中使用的术语“和/或”仅仅是一种描述关联对象的关联关系，表示可以存在三种关系，例如，A 和/或 B，可以表示：单独存在 A，同时存在 A 和 B，单独存在 B 这三种情况。另外，本文中字符“/”，一般表示前后关联对象是一种“或”的关系。

应当理解，尽管在本申请实施例中可能采用术语第一、第二、第三等来描述，但这些描述不应限于这些术语。这些术语仅用来将描述区分开。例如，在不脱离本申请实施例范围的情况下，第一也可以被称为第二，类似地，第二也可以被称为第一。

取决于语境，如在此所使用的词语“如果”、“若”可以被解释成为“在……时”或“当……时”或“响应于确定”或“响应于检测”。类似地，取决于语境，短语“如果确定”或“如果检测（陈述的条件或事件）”可以被解释成为“当确定时”或“响应于确定”或“当检测（陈述的条件或事件）时”或“响应于检测（陈述的条件或事件）”。

还需要说明的是，术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含，从而使得包括一系列要素的商品或者装置不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种商品或者装置所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括所述要素的商品或者装置中还存在另外的相同要素。

特别需要说明的是，在说明书中存在的符号和/或数字，如果在附图说明中未被标记的，均不是附图标记。

参照图 1，一种日志敏感信息检测方法，包括：

S101，响应于敏感信息检测信号，基于定义的语言规则，获取源代码中包含的命名变量信息以及程序调用接口函数信息；

具体的，在此步骤中首先定义语言规则，根据定义的语言规则在源代码中获取命名变量信息以及程序调用接口函数信息；

可以理解的是，此步骤不依赖日志，从源代码中进行命名变量信息以及程序调用接口函数信息的获取，定义的语言规则可以针对不同的数据库以及不同的编程语言进行针对性定义，目的是在不同种类的数据库以及不同种类的编程语言中获取命名变量信息以及程序调用接口函数信息。

在其中一些应用中，响应于敏感信息检测信号，基于定义的语言规则，获取源代码中包含的命名变量信息以及程序调用接口函数（API 函数）信息，具体包括所述定义的语言规则包括，代码编程规则；基于所述编程规则，获取源代码中包含的命名变量信息以及程序调用接口函数信息；其中，所述命名变量信息包括，数据库服务器地址（database_server）、数据名称（database_name）、用户名（database_user）以及用户密码（database_password）；

所述调用接口函数信息包括，连接数据操作（connect_database）、登录系统（login_system）以及打开服务（open_service）。

可以理解的是，在此应用中，根据符合相应数据库中的编程代码规则，相应的获取源代码中的命名变量信息以及程序调用接口函数信息，例如命名变量信息可以是 database_server、database_name、database_user 以及 database_password，调用接口函数信息可以是 connect_database、login_system 以及 open_service，上述命名变量信息以及程序调用接口函数信息都是有可能涉及敏感信息的关键位置，因此，从源代码中将相关信息进行获取。

S102，基于定义的检索规则，确定所述命名变量信息以及所述程序调用接口函数信息在所述源代码中的位置信息；

具体的，在此步骤中，首先定义检索规则，根据定义的检索规则在源代码中确认命名变量信息以及所述程序调用接口函数信息的位置；

可以理解的是，此步骤的目的是从源代码中对命名变量信息以及所述程序调用接口函数信息进行快速定位，检索规则可以进行针对性设置，可以通过正则表达式进行匹配。

在其中一些应用中，基于定义的检索规则，确定所述命名变量信息以及所述程序调用接口函数信息在所述源代码中的位置信息，具体包括所述定义的检索规则包括预置关键字列表，对所述命名变量信息进行检索；对所述源代码进行扫描，对所述程序调用接口函数信息进行检索；基于检索到的所述命名变量信息以及所述程序调用接口函数信息，确定二者在所述源代码中的所述位置信息；

可以理解的是，在此应用中，通过预设关键词列表对命名变量信息进行检索，关键词列表可以根据想要获取的命名变量进行灵活摄者，由于程序调用接口函数信息通常在程序中存在较少的表达方式，因此对源代码进行扫描即可进行获取，最后根据检索或扫描的命名变量信息以及所述程序调用接口函数信息，确认二者相应的位置信息。

S103，基于所述位置信息，对附加所述命名变量信息以及所述程序调用接口函数信息的日志输出代码进行敏感信息检测；

具体的，在此步骤中，根据命名变量信息以及程序调用接口函数信息在源代码中的位置信息，对附加命名变量信息以及程序调用接口函数信息的日志输出代码进行敏感信息检测；

可以理解的是，此步骤中，当确定命名变量信息以及程序调用接口函数信息在源代码中的位置信息，即可对相应的日志进行定位，根据日志的输出代码进行敏感信息检测。

在其中一些应用中，基于所述位置信息，对附加所述命名变量信息以及所述程序调用接口函数信息的日志输出代码进行敏感信息检测，具体包括基于确定的所述命名变量信息以及所述程序调用接口函数信息在所述源代码中的位置信息，确定与之相关联的所述日志；对所述日志的输出代码进行敏感信息检测，确认所述日志输出代码是否包含所述敏感信息；

可以理解的是，在此应用中，根据命名变量信息以及所述调用接口函数信息在源代码中所在的位置，对相关联的日志进行查询确认，确认日志输出代码是否包含所述敏感信息。

在其中一些实施例中，为了使敏感信息检测的更为准确，基于所述位置信息，对附加所述命名变量信息以及所述程序调用接口函数信息的日志输出代码进行敏感信息检测后，所述方法还包括基于所述日志输出代码，对获取的所述命名变量信息以及所述程序调用接口函数信息进行验证，确认所述日志中是否检测到所述敏感信息；

可以理解的是，在此实施例中，通过对命名变量信息以及调用接口函数信息在源代码中所在的位置相关联日志输出代码，与获取的命名变量信息以及程序调用接口函数信息进行一一验证的分析处理，可进一步保证敏感信息检测结果的准确性。

S104，当检测出所述敏感信息时，对所述敏感信息进行风险提示。

可以理解的是，在此步骤中，需要对敏感信息进行风险提示；

在其中一些应用中，当检测到所述敏感信息时，对所述敏感信息进行标记；基于所述标记，进行相应的风险提示。

可以理解的是，在此应用中，针对不同的敏感信息可以划分相应的风险等级，根据不同的风险等级定义不同的标记，例如 1 级 2 级 3 级等，然后根据不同的风险等级设置不同的风险提示，风险提示可以是文字形式的，可以写明相应的风险说明；

进一步的，对所述敏感信息进行风险提示后，所述方法还包括对标记的所述敏感信息进行脱敏处理。

在其中一些应用中，此处的脱敏处理，可以是对检测到的敏感信息进行删除；

可以理解的是，通过上述方法可以不依赖日志的输出形式，可以检测出涉及敏感信息位置，并进一步检测是否存续敏感信息泄漏的风险，检测到风险的同时，可明确问题所在源代码位置，便于确认和修复问题，不依赖于程序的动态运行，检测结果更准确、全面且效率更高。

下面将结合图 2 说明本发明日志敏感信息检测方法在一些应用中的实施例，如图 2 所示：

为了更好的说明本专利方案，以数据库连接的场景为例进行说明。

一个典型数据库连接操作，首先设置数据库用户名和密码变量，调用数据库连接 API 函数连接数据库服务。

一、在代码中可能存在如下命名的变量（不同数据库和编程语言有不同命名，基本上都是相关含义的英文单词）：

database_server，数据库服务器地址；

database_name，数据名称；

database_user，用户名；

database_password，用户密码；

调用的数据 api 函数可能命名（不同数据库和编程语言有不同命名，基本上都是相关含义的英文单词）：

connect_database()，连接数据操作；

二、结合源码，通过变量命名和程序调用接口函数（API 函数），可以捕捉到可能的敏感信息。

1、识别敏感变量：程序源代码中，变量命名通常有一定的规范，例如上面提到的 database_user 和 database_password 会直观表达所代表的含义。通过一些预置的关键字列表或正则表达式规则进行匹配，可以检索到此类名称的变量。

2、识别敏感函数调用：代码中调用的 API 函数，通常也是有规范的命名方式，而且更严格和正规，例如上述的 connect_database（连接数据库），以及一些常用的 API 命名包括 login_system（登录系统）、open_service（打开某个服务，如 web 应用、ftp）等，这类接口通常都要求用户名、密码等信息。

3、系统的 API 函数在编码时是确定的；并且相比代码总量，涉及敏感信息的 API 函数是少量的。完全可以通过扫描代码，找到所有调用此类函数的代码。

这里是用用户名、密码进行举例，采用类似逻辑，可检索到其他敏感信息的代码位置。其他敏感信息可能包括：硬件设备 ID，用户个人信息，金融相关信息等。

找到上述涉及敏感信息源码，算法扫描其附日志输出代码，并检测日志输出中是否包含敏感信息。如果包含敏感信息，则进行标记或者提示风险。

为了提升准确性，还可对程序运行的日志输出进行分析处理，和源码的扫描结果进行验证，确认日志中是否出现检测到的敏感信息。这一步是为了提升准确性，核心算法还是上述的操作步骤。

本方法对日志输出形式不敏感，可以更加准确的识别各种日志中敏感信息。不同的日志输出方式，会影响信息的识别。

例如：在日志文件中有如下 2 条日志信息，开发人员为了调试打印出来用户名（beijing）和密码（332211）。

A：connect Beijing_market_database with user=beijing and password=332211

B：beijing-332211

不同的开发人员有不同的习惯打印日志。日志 A 的信息更加明确，日志 B 的信息比较模糊，如果采用事后日志分析的方式识别敏感信息，日志 A 容易识别，日志 B 很可能会被遗漏。

本方法不依赖日志输出内容，而是根据代码信息进行筛选，可以更加有效、准确的识别敏感信息。

通过上述方法，可实现：

基于源代码自动检测敏感信息。

对 log 内容和格式不敏感，这种检测方法不依赖日志的输出形式，而且可以检测出涉及敏感信息位置，并进一步检测是否存续敏感信息泄漏的风险。

检测到风险的同时，可明确问题所在源代码位置，便于确认和修复问题。

不依赖于程序的动态运行，检测结果更准确、全面。一个软件系统会有多种逻辑分支，动态运行时无法人为控制系统实际运行的逻辑分支，有的分支可能不会执行到，尤其对于复杂系统。相关日志输出代码就不会被执行，也就不会执行泄露信息的代码，事后分析日志的方法无法检测到这种情况。通过源码检测可以避免这些问题，可以对源码进行完整的扫描，不会有遗漏。

效率更高。系统运行时，同样的代码逻辑会重复执行，日志会重复打印（对应代码每次执行都会打印日志），日志量会非常大。可能需要开发分析工具软件，并有较大的分析工作量。使用源码分析方法就会高效很多，源代码不会有重复，检测量比事后日志分析少的多。

对于上述实施例公开的方法步骤，出于简单描述的目的将方法步骤表述为一系列的动作组合，但是本领域技术人员应该知悉，本发明实施例并不受所描述的动作顺序的限制，因为依据本发明实施例，某些步骤可以采用其他顺序或者同时进行。其次，本领域技术人员也应该知悉，说明书中所描述的实施例均属于优选实施例，所涉及的动作并不一定是本发明实施例所必须的。

如图 3 所示，本发明还提供一种日志敏感信息检测系统，包括：

信息获取模块 201，配置为响应于敏感信息检测信号，基于定义的语言规则，获取源代码中包含的命名变量信息以及程序调用接口函数信息；

位置信息确认模块 202，配置为基于定义的检索规则，确定所述命名变量信息以及所述程序调用接口函数信息在所述源代码中的位置信息；

敏感信息检测模块 203，配置为基于所述位置信息，对附加所述命名变量信息以及所述程序调用接口函数信息的日志输出代码进行敏感信息检测；

敏感信息提示模块 204，配置为当检测出所述敏感信息时，对所述敏感信息进行风险提示。

值得注意的是，虽然在本发明实施例中只披露了一些基本功能模块，但并不意味着本系统的组成仅仅局限于上述基本功能模块，相反，本实施例所要表达的意思是：在上述基本功能模块的基础之上本领域技术人员可以结合现有技术任意添加一个或多个功能模块，形成无穷多个实施例或技术方案，也就是说本系统是开放式而非封闭式的，不能因为本实施例仅仅披露了个别基本功能模块，就认为本发明权利要求的保护范围局限于所公开的基本功能模块。同时，为了描述的方便，描述以上装置时以功能分为各种单元、模块分别描述。当然

在实施本发明时可以把各单元、模块的功能在同一个或多个软件和/或硬件中实现。

如图 4 所示，本发明还提供一种电子设备，包括：处理器、通信接口、存储器和通信总线，其中，处理器，通信接口，存储器通过通信总线完成相互间的通信；所述存储器中存储有计算机程序，当所述计算机程序被所述处理器执行时，使得所述处理器执行日志敏感信息检测方法的步骤。

图 4 是本发明实施例提供的一种电子设备的结构示意图。如图 4 所示结构，本发明实施例中提供的电子设备包括：一个或多个处理器 710 和存储装置 720；该电子设备中的处理器 710 可以是一个或多个，图 4 中以一个处理器 710 为例；存储装置 720 用于存储一个或多个程序；所述一个或多个程序被所述一个或多个处理器 710 执行，使得所述一个或多个处理器 710 实现如本发明实施例中任一项所述的日志敏感信息检测方法。

该电子设备还可以包括：输入装置 730 和输出装置 740。

该电子设备中的处理器 710、存储装置 720、输入装置 730 和输出装置 740 可以通过总线或其他方式连接，图 4 中以通过总线连接为例。

该电子设备中的存储装置 720 作为一种计算机可读存储介质，可用于存储一个或多个程序，所述程序可以是软件程序、计算机可执行程序以及模块，如本发明实施例中所提供的日志敏感信息检测方法对应的程序指令/模块。处理器 710 通过运行存储在存储装置 720 中的软件程序、指令以及模块，从而执行电子设备的各种功能应用以及数据处理，即实现上述方法实施例中日志敏感信息检测方法。

存储装置 720 可包括存储程序区和存储数据区，其中，存储程序区可存储操作系统、至少一个功能所需的应用程序；存储数据区可存储根据电子设备的使用所创建的数据等。此外，存储装置 720 可以包括高速随机存取存储器，还可以包括非易失性存储器，例如至少一个磁盘存储器件、闪存器件、或其他非

易失性固态存储器件。在一些实例中，存储装置 720 可进一步包括相对于处理器 710 远程设置的存储器，这些远程存储器可以通过网络连接至设备。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

输入装置 730 可用于接收输入的数字或字符信息，以及产生与电子设备的用户设置以及功能控制有关的键信号输入。输出装置 740 可包括显示屏等显示设备。

本发明还提供一种计算机可读存储介质，其存储有可由电子设备执行的计算机程序，当所述计算机程序在所述电子设备上运行时，使得所述电子设备执行日志敏感信息检测方法的步骤。

具体的，本发明实施例的计算机存储介质，可以采用一个或多个计算机可读的介质的任意组合。计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质。计算机可读存储介质例如可以是——但不限于——电、磁、光、电磁、红外线、或半导体的系统、装置或器件，或者任意以上的组合。计算机可读存储介质的更具体的例子(非穷举的列表)包括：具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM 或闪存)、光纤、便携式紧凑磁盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本实施例中，计算机可读存储介质可以是任何包含或存储程序的有形介质，该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。

最后应说明的是：以上各实施例仅用以说明本发明的技术方案，而非对其限制；尽管参照前述各实施例对本发明进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分或者全部技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本发明各实施例技术方案的范围。

权 利 要 求 书

1、一种日志敏感信息检测方法，其特征在于，包括：

响应于敏感信息检测信号，基于定义的语言规则，获取源代码中包含的命名变量信息以及程序调用接口函数信息；

基于定义的检索规则，确定所述命名变量信息以及所述程序调用接口函数信息在所述源代码中的位置信息；

基于所述位置信息，对附加所述命名变量信息以及所述程序调用接口函数信息的日志输出代码进行敏感信息检测；

当检测出所述敏感信息时，对所述敏感信息进行风险提示。

2、根据权利要求 1 所述的日志敏感信息检测方法，其特征在于，基于所述位置信息，对附加所述命名变量信息以及所述程序调用接口函数信息的日志输出代码进行敏感信息检测后，所述方法还包括：

基于所述日志输出代码，对获取的所述命名变量信息以及所述程序调用接口函数信息进行验证，确认所述日志中是否检测到所述敏感信息。

3、根据权利要求 1 所述的日志敏感信息检测方法，其特征在于，响应于敏感信息检测信号，基于定义的语言规则，获取源代码中包含的命名变量信息以及程序调用接口函数信息，具体包括：

所述定义的语言规则包括，代码编程规则；

基于所述编程规则，获取源代码中包含的命名变量信息以及程序调用接口函数信息；

其中，所述命名变量信息包括，数据库服务器地址、数据名称、用户名以及用户密码；

所述调用接口函数信息包括，连接数据操作、登录系统以及打开服务。

4、根据权利要求 1 所述的日志敏感信息检测方法，其特征在于，基于定义的检索规则，确定所述命名变量信息以及所述程序调用接口函数信息在所述源代码中的位置信息，具体包括：

所述定义的检索规则包括：

预置关键字列表，对所述命名变量信息进行检索；

对所述源代码进行扫描，对所述程序调用接口函数信息进行检索；

基于检索到的所述命名变量信息以及所述程序调用接口函数信息，确定二者在所述源代码中的所述位置信息。

5、根据权利要求 1 所述的日志敏感信息检测方法，其特征在于，基于所述位置信息，对附加所述命名变量信息以及所述程序调用接口函数信息的日志输出代码进行敏感信息检测，具体包括：

基于确定的所述命名变量信息以及所述程序调用接口函数信息在所述源代码中的位置信息，确定与之相关联的所述日志；

对所述日志的输出代码进行敏感信息检测，确认所述日志输出代码是否包含所述敏感信息。

6、根据权利要求 1 所述的日志敏感信息检测方法，其特征在于，当检测出所述敏感信息时，对所述敏感信息进行风险提示，具体包括：

当检测到所述敏感信息时，对所述敏感信息进行标记；

基于所述标记，进行相应的风险提示。

7、根据权利要求 6 所述的日志敏感信息检测方法，其特征在于，对所述敏感信息进行风险提示后，所述方法还包括：

对标记的所述敏感信息进行脱敏处理。

8、一种日志敏感信息检测系统，其特征在于，包括：

信息获取模块，配置为响应于敏感信息检测信号，基于定义的语言规则，获取源代码中包含的命名变量信息以及程序调用接口函数信息；

位置信息确认模块，配置为基于定义的检索规则，确定所述命名变量信息以及所述程序调用接口函数信息在所述源代码中的位置信息；

敏感信息检测模块，配置为基于所述位置信息，对附加所述命名变量信息以及所述程序调用接口函数信息的日志输出代码进行敏感信息检测；

敏感信息提示模块，配置为当检测出所述敏感信息时，对所述敏感信息进行风险提示。

9、一种电子设备，其特征在于，包括：处理器、通信接口、存储器和通信总线，其中，处理器，通信接口，存储器通过通信总线完成相互间的通信；所述存储器中存储有计算机程序，当所述计算机程序被所述处理器执行时，使得所述处理器执行权利要求 1 至 7 中任一项所述方法的步骤。

10、一种计算机可读存储介质，其特征在于，其存储有可由电子设备执行的计算机程序，当所述计算机程序在所述电子设备上运行时，使得所述电子设备执行权利要求 1 至 7 中任一项所述方法的步骤。

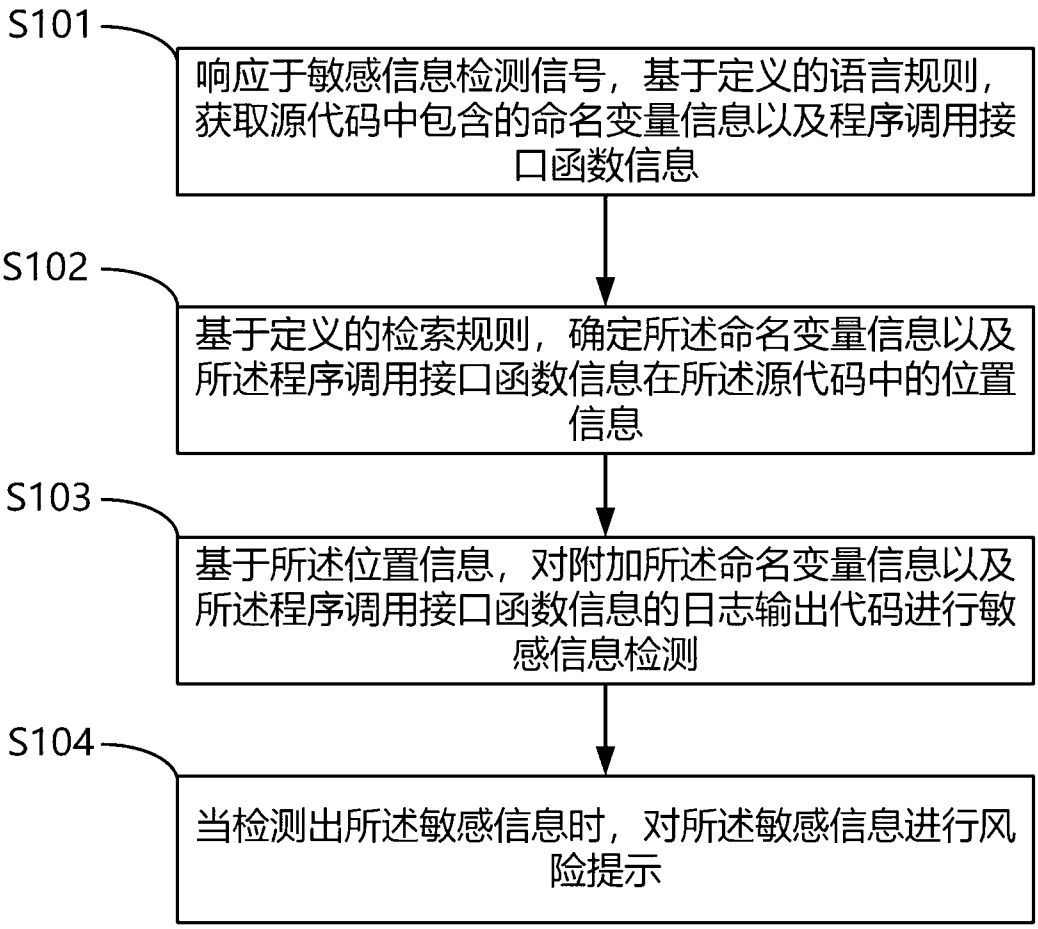

S101
响应于敏感信息检测信号，基于定义的语言规则，获取源代码中包含的命名变量信息以及程序调用接口函数信息
S102
基于定义的检索规则，确定所述命名变量信息以及所述程序调用接口函数信息在所述源代码中的位置信息
S103
基于所述位置信息，对附加所述命名变量信息以及所述程序调用接口函数信息的日志输出代码进行敏感信息检测
S104
当检测出所述敏感信息时，对所述敏感信息进行风险提示

图 1

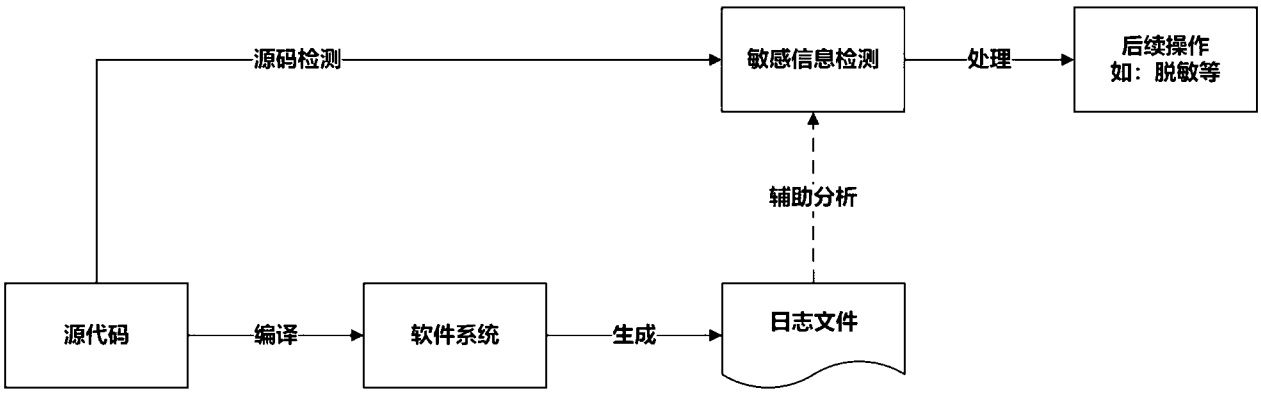


图 2

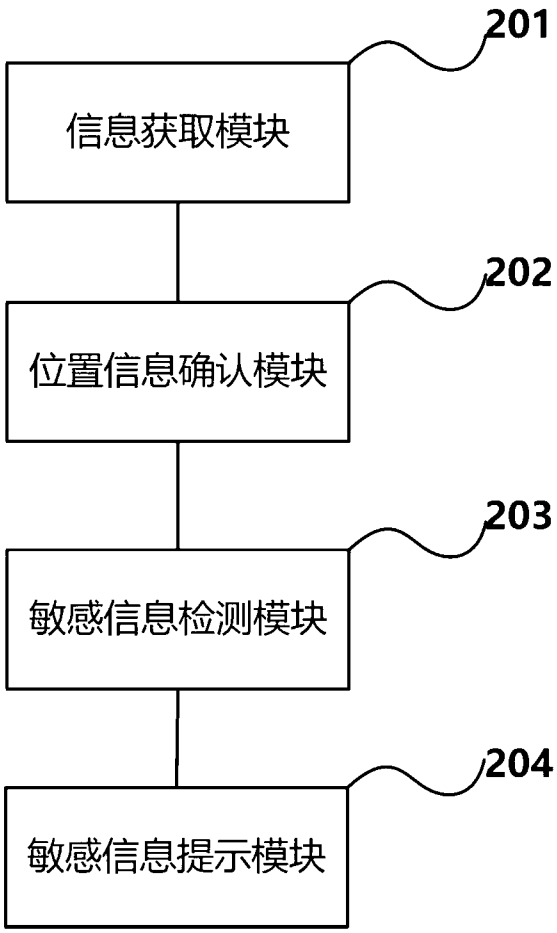


图 3

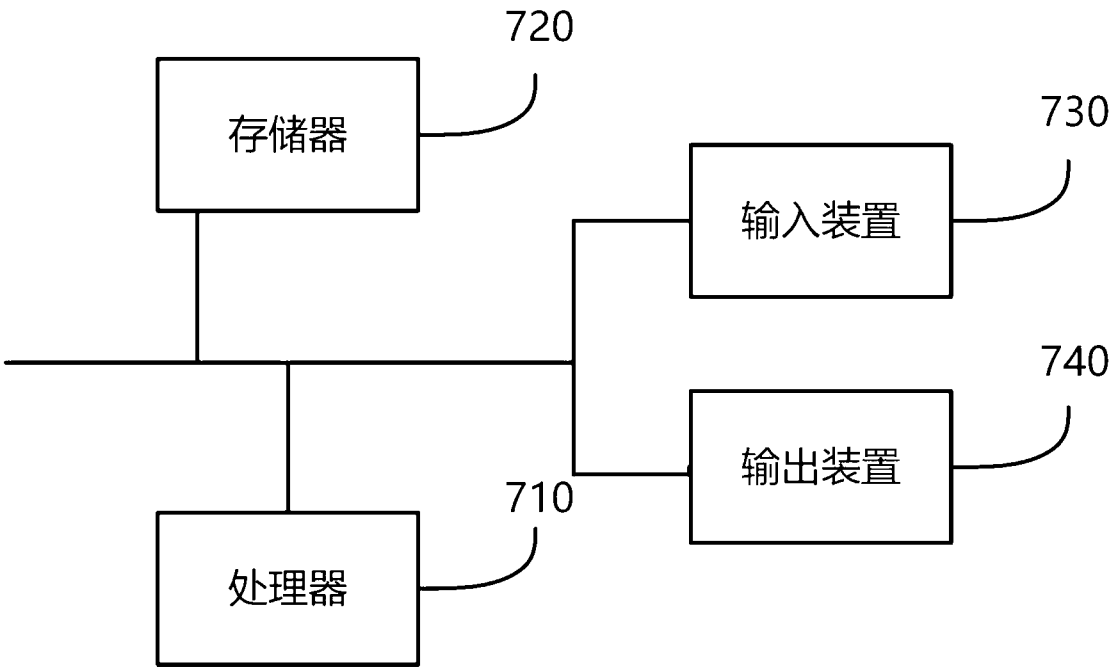


图 4

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2024/112983

A. CLASSIFICATION OF SUBJECT MATTER

G06F 21/62(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)

IPC: 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)

VEN, CNABS, CNTXT, WOTXT, EPTXT, USTXT, CNKI, IEEE: 日志, 源代码, 搜索, 检索, 查询, 扫描, 检测, 变量, 调用, 接口, 数据库, 连接, 用户名, 密码, 密钥, 登录, 打印, 输出, 敏感, 隐私, 脱敏, log, program, language, variable, query, search, scan, API, interface, call, database, connection, username, password, key, login, print, output, privacy, sensitive

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 117910030 A (CHINA FAW GROUP CORP., LTD. et al.) 19 April 2024 (2024-04-19) claims 1-10	1-10
A	CN 110598411 A (TENCENT TECHNOLOGY (SHENZHEN) CO., LTD.) 20 December 2019 (2019-12-20) description, paragraphs [0029]-[0119], and figures 1-8	1-10
A	CN 109614814 A (PING AN PUHUI ENTERPRISE MANAGEMENT CO., LTD.) 12 April 2019 (2019-04-12) entire document	1-10
A	CN 112035354 A (BEIJING ZHIZHANGYI SCIENCE & TECHNOLOGY CO., LTD.) 04 December 2020 (2020-12-04) entire document	1-10
A	US 2020285772 A1 (INTERNATIONAL BUSINESS MACHINES CORP.) 10 September 2020 (2020-09-10) entire document	1-10

☑ 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
"D" document cited by the applicant in the international application
"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
05 November 2024	**12 November 2024**
Name and mailing address of the ISA/CN **China National Intellectual Property Administration (ISA/CN)** **China No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088**	Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2024/112983

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2022327016 A1 (EMC IP HOLDING COMPANY LLC) 13 October 2022 (2022-10-13) entire document	1-10

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2024/112983

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 117910030 A	19 April 2024	None	
CN 110598411 A	20 December 2019	None	
CN 109614814 A	12 April 2019	None	
CN 112035354 A	04 December 2020	None	
US 2020285772 A1	10 September 2020	US 2020285773 A1	10 September 2020
US 2022327016 A1	13 October 2022	CN 115202995 A	18 October 2022

国际检索报告

国际申请号
PCT/CN2024/112983

A. 主题的分类

G06F 21/62(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)

IPC: 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））

VEN，CNABS，CNTXT，WOTXT，EPTXT，USTXT，CNKI，IEEE：日志，源代码，搜索，检索，查询，扫描，检测，变量，调用，接口，数据库，连接，用户名，密码，密钥，登录，打印，输出，敏感，隐私，脱敏，log，program，language，variable，query，search，scan，API，interface，call，database，connection，username，password，key，login，print，output，privacy，sensitive

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 117910030 A (中国第一汽车股份有限公司 等) 2024年4月19日 (2024 - 04 - 19) 权利要求1-10	1-10
A	CN 110598411 A (腾讯科技（深圳）有限公司) 2019年12月20日 (2019 - 12 - 20) 说明书第[0029]-[0119]段、图1-8	1-10
A	CN 109614814 A (平安普惠企业管理有限公司) 2019年4月12日 (2019 - 04 - 12) 全文	1-10
A	CN 112035354 A (北京指掌易科技有限公司) 2020年12月4日 (2020 - 12 - 04) 全文	1-10
A	US 2020285772 A1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 2020年9月10日 (2020 - 09 - 10) 全文	1-10
A	US 2022327016 A1 (EMC IP HOLDING COMPANY LLC) 2022年10月13日 (2022 - 10 - 13) 全文	1-10

☐ 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

“D” 申请人在国际申请中引证的文件

“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
2024年11月5日	2024年11月12日
ISA/CN的名称和邮寄地址 中国国家知识产权局 中国北京市海淀区蓟门桥西土城路6号 100088	受权官员 董洪梅 电话号码 (+86) 010-53961528

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2024/112983

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 117910030 A	2024年4月19日	无	
CN 110598411 A	2019年12月20日	无	
CN 109614814 A	2019年4月12日	无	
CN 112035354 A	2020年12月4日	无	
US 2020285772 A1	2020年9月10日	US 2020285773 A1	2020年9月10日
US 2022327016 A1	2022年10月13日	CN 115202995 A	2022年10月18日