

(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)



(10) 国际公布号
WO 2025/123745 A1

- (51) 国际专利分类号:
G06F 16/18 (2019.01) **G06F 16/13** (2019.01)
G06F 16/17 (2019.01)
- (21) 国际申请号: PCT/CN2024/112984
- (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:
202311709368.0 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)。蒲军霖 (**PU, Junlin**); 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 FILE STORAGE METHOD AND SYSTEM, ELECTRONIC DEVICE, AND STORAGE MEDIUM

(54) 发明名称: 一种日志文件存储方法、系统、电子设备及存储介质

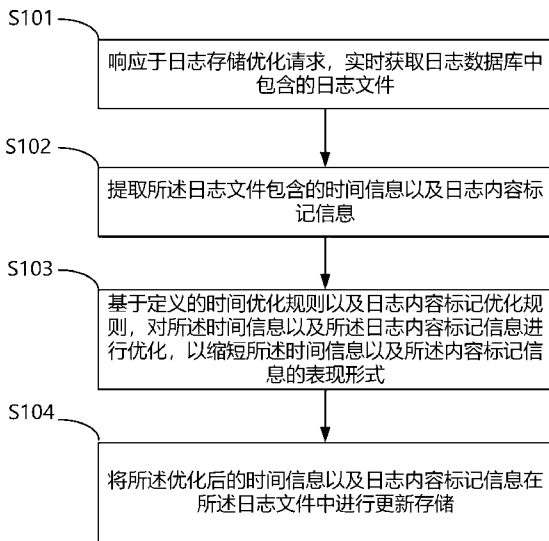


图 1

- S101 In response to a log storage optimization request, obtain log files contained in a log database in real time
- S102 Extract time information and log content tag information contained in each log file
- S103 Optimize the time information and the log content tag information on the basis of a defined time optimization rule and a defined log content tag optimization rule, so as to shorten expression forms of the time information and the content tag information
- S104 Update and store the optimized time information and the optimized log content tag information in the corresponding log file

(57) **Abstract:** Disclosed in the present invention are a log file storage method and system, an electronic device, and a storage medium. The method comprises: in response to a log storage optimization request, obtaining log files contained in a log database in real time; extracting time information and log content tag information contained in each log file; optimizing the time information and the log content tag information on the basis of a defined time optimization rule and a defined log content tag optimization rule, so as to shorten expression forms of the time information and the content tag information; and updating and storing the optimized time information and the optimized log content tag information in the corresponding log file. The method can improve the utilization rate of a log storage space, enabling preservation of more effective logs in the same space, correspondingly increasing the log preservation duration, and facilitating subsequent system development and debugging and subsequent vehicle operation.



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))。

(57) 摘要: 本发明公开了一种日志文件存储方法、系统、电子设备及存储介质, 包括响应于日志存储优化请求, 实时获取日志数据库中包含的日志文件; 提取所述日志文件包含的时间信息以及日志内容标记信息; 基于定义的时间优化规则以及日志内容标记优化规则, 对所述时间信息以及所述日志内容标记信息进行优化, 以缩短所述时间信息以及所述内容标记信息的表现形式; 将所述优化后的时间信息以及日志内容标记信息在所述日志文件中进行更新存储。通过上述方法可提高日志存储空间利用率, 同样空间保存更多有效日志, 保存日志时长也相应的增加, 有利于后续的系统开发调试以及后续的车辆运营。

一种日志文件存储方法、系统、电子设备及存储介质

技术领域

本发明涉及存储技术领域，特别是涉及一种日志文件存储方法、系统、电子设备及存储介质。

5 背景技术

日志系统是车机系统上的关键基础子系统。车机系统不同于其他系统开发，比如 PC 应用开发、Web 应用开发等可以随时进行联机调试，车机系统只有在早期开发阶段以实验室开发为主，后续的开发和运营都是和开发环境脱离的，比如：路测、用户实际驾驶等场景，都无法进行联调，这就导致遇到问题时难以定位问题或还原问题现场。这种情况下系统和应用日志可以提供关键信息辅助定位问题、还原现象。如果离开日志，不只开发效率会极大降低，甚至整个开发难以进行。

日志系统是底层模块，最终用户不会直接使用，系统资源使用比较有限。日志系统是整个车机系统的一部分，在运行过程中会占用 CPU、内存、存储空间等系统资源，同时由于日志系统是“非功能模块”，用户不会直接感知，而是为其他系统服务或应用提供支持，这就要求日志模块尽量占用少的系统资源，或者在有限的系统资源下，可以更加高效的工作。

现有技术中，普遍采用的方案是为日志文件限定个数和单个文件的大小上限，即限制日志文件占用空间的总大小。当日志内容超出限制是，就会覆盖日志文件中最旧的一个日志文件，即循环覆盖策略。

CN111538704A 的专利文件公开了一种日志优化方法、装置、设备及存储介质，该日志优化方法通过提取日志文件集中每条日志的属性特征和属性值，对其进行归类，得到属性集合，构建树状图，然后选取目标属性特征，根据预设的抽象摘要算法对目标属性特征下的属性值进行抽象处理，获得该日志文件集的关键摘要，并以所述关键摘要代替当前系统所存储的所述日志文件集合，最后分析关键摘要可以提高异常诊断的效率。

上述现有技术的方案中存在优化过程繁琐，优化效率较低，不利于系统的开发，还会带来如下问题：

1、给定的日志存储空间，保存更少的日志时长，不利定位和解决问题。

2、在后续运营中，需要远程抓取车辆日志，还会涉及到日志的网络传输、
5 云端存储、内容自动化分析等。存储低效的日志（重复冗余或无效日志内容），会有占用更大的带宽、更长传输时间、更多的出错概率、更大的云端存储、以及让自动分析效率更低。

3、总体上增大研发、运营的成本，而且会日积月累，不断增加。

4、如果压缩算法动态处理日志，不仅占用资源高（CPU，内存等），而
10 且也无法直接查看日志内容。

因此，如何提供一种日志文件存储方法，提高日志优化效率以及优化效果，是本领域技术人员有待解决的技术问题。

发明内容

为解决现有技术中日志文件存储优化效率较低以及优化效果较差的技术
15 问题，本发明提供一种日志文件存储方法、系统、电子设备及存储介质。

为实现本发明目的提供的一种日志文件存储方法，包括：

响应于日志存储优化请求，实时获取日志数据库中包含的日志文件；

提取所述日志文件包含的时间信息以及日志内容标记信息；

基于定义的时间优化规则以及日志内容标记优化规则，对所述时间信息以
20 及所述日志内容标记信息进行优化，以缩短所述时间信息以及所述内容标记信息的表现形式；

将所述优化后的时间信息以及日志内容标记信息在所述日志文件中进行更新存储。

在其中一些具体实施例中，所述定义的时间优化规则，具体包括：

25 基于所述时间信息之间的时间差值，对所述时间信息进行优化。

在其中一些具体实施例中，所述定义的日志内容标记优化规则，具体包括：

建立对应所述日志内容标记的映射关系值；

基于所述映射关系值，对所述日志内容标记进行优化。

在其中一些具体实施例中，基于定义的时间优化规则，对所述时间信息进行优化，以缩短所述时间信息的表现形式，具体包括：

5 判断当前日志是否为所述日志文件的第一条记录；

当所述日志不是所述日志文件的第一条记录时，计算当前所述日志的所述时间信息与上一条日志的所述时间信息的差值；

依据所述时间信息的差值，对所述当前日志的所述时间信息进行修改，以缩短所述时间信息的表现形式。

10 在其中一些具体实施例中，当所述日志是所述日志文件的第一条记录时，不对所述日志的时间信息进行优化，保留完整的所述时间信息；

计算所述日志的所述时间信息与下一条日志的所述时间信息的差值；

依据所述时间信息的差值，对所述下一条的所述时间信息进行修改，以缩短所述时间信息的表现形式。

15 在其中一些具体实施例中，基于定义的日志内容标记优化规则，对所述日志内容标记信息进行优化，以缩短所述内容标记信息的表现形式，具体包括：

建立对应所述日志内容标记的映射关系值；

在所述日志数据库中依次保存与所述日志内容标记对应的映射关系值；

20 依据所述各映射关系值，修改与所述各映射关系值对应的各所述日志内容标记，以缩短所述日志内容标记的表现形式。

在其中一些具体实施例中，依据所述各映射关系值，修改与所述各映射关系值对应的各所述日志内容标记，以缩短所述日志内容标记的表现形式后，所述方法还包括：

25 当接收到新日志时，判断是否存在与所述新日志的日志内容标记对应的所述映射关系值；

当存在所述映射关系值时，修改所述新日志的日志内容标记，以缩短所述日志内容标记的表现形式；

当不存在所述映射关系值时，建立与所述新日志的日志内容标记对应的所述映射关系值，并依据建立的所述映射关系值修改所述新日志的日志内容标记，

5 以缩短所述日志内容标记的表现形式。

基于同一构思，本发明还提供一种日志文件存储系统，包括：

日志文件获取模块，配置为响应于日志存储优化请求，实时获取日志数据库中包含的日志文件；

10 日志文件提取模块，配置为提取所述日志文件包含的时间信息以及日志内容标记信息；

日志文件优化模块，配置为基于定义的时间优化规则以及日志内容标记优化规则，对所述时间信息以及所述日志内容标记信息进行优化，以缩短所述时间信息以及所述内容标记信息的表现形式；

15 日志文件存储模块，配置为将所述优化后的时间信息以及日志内容标记信息在所述日志文件中进行更新存储。

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文件存储方法的步骤。

20 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文件存储方法的步骤。

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文件存储方法、系统、电子设备及存储介质，可提高日志存储空间利用率，同样空间保存更多有效日志，保存日志时长也相应的增加，有利于后续的系统开发调试以及后续的车辆运营。

附图说明

- 5 图 1 是本发明一种日志文件存储方法在一些具体实施例的流程示意图；
图 2 是本发明一种日志文件存储方法在一些应用中的流程示意图；
图 3 是本发明一种日志文件存储方法在另一些应用中的流程示意图；
图 4 是本发明一种日志文件存储系统在一些具体实施例的结构示意图；
图 5 是本发明一种电子设备在一些具体实施例的结构示意图。

10 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。

- 15 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
20 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 1，一种日志文件存储方法，包括：

S101，响应于日志存储优化请求，实时获取日志数据库中包含的日志文件；

具体的，在此步骤中，当接收到日志存储优化请求时，实时获取日志数据库中包含的日志文件；

可以理解的是，日志数据库中包含了多个日志文件，同时日志文件也在源源不断的进入到日志数据库中，此步骤当接收到日志存储优化请求时，对各日志文件以及新进入的日志文件进行实时获取，日志存储优化请求也可更改为常开状态，这样日志数据库将实时进行优化。

S102，提取所述日志文件包含的时间信息以及日志内容标记信息；

具体的，在步骤中，在获取的各日志文件中，提取各日志文件包含的时间信息以及日志内容标记信息；

可以理解的是，日志文件包含的时间信息，可理解为日志文件中每一条日志记录的时间字符串，日志内容标记信息，可理解为日志文件中每一条日志记录的日志 tag，从获取的日志文件中，提取其中包含的时间字符串以及日志 tag。

5 S103，基于定义的时间优化规则以及日志内容标记优化规则，对所述时间信息以及所述日志内容标记信息进行优化，以缩短所述时间信息以及所述内容标记信息的表现形式；

具体的，在此步骤中，首先定义时间优化规则以及日志内容标记优化规则，接着根据定义的时间优化规则以及日志内容标记优化规则，对时间信息以及日志内容标记信息进行优化，以缩短时间信息以及内容标记信息的表现形式；

10 可以理解的是，定义的时间优化规则以及日志内容标记优化规则，对应优化的是日志文件中包含的时间信息以及内容标记信息，优化的目的是缩短时间信息以及内容标记信息的表现形式，可理解为缩短时间字符串以及日志 tag。

在其中一些应用中，所述定义的时间优化规则，具体包括基于所述时间信息之间的时间差值，对所述时间信息进行优化；

15 可以理解的是，在此应用中，将各时间信息之间的是时间差定义为时间优化规则，可通过时间差替换相应的时间字符串。

进一步的，在此应用中，基于定义的时间优化规则，对所述时间信息进行优化，以缩短所述时间信息的表现形式，具体包括判断当前日志是否为所述日志文件的第一条记录；当所述日志不是所述日志文件的第一条记录时，计算当前所述日志的所述时间信息与上一条日志的所述时间信息的差值；依据所述时间信息的差值，对所述当前日志的所述时间信息进行修改，以缩短所述时间信息的表现形式；

20

可以理解的是，在此应用中首先判断当前的日志是否是日志文件的第一条记录，当日志不是日志文件的第一条记录时，表示至少还存在此记录的上一条记录，计算这两条记录的时间信息差值，即时间字符串的时间差值，通过时间

25

字符串的时间差值，修改此条记录的时间信息，即修改此条记录的时间字符串为时间差值，以此缩短时间信息的表现形式，优化了存储空间，持续此环节则可节约大量的存储空间。

进一步的，在此应用中，当所述日志是所述日志文件的第一条记录时，不对所述日志的时间信息进行优化，保留完整的所述时间信息；计算所述日志的所述时间信息与下一条日志的所述时间信息的差值；依据所述时间信息的差值，对所述下一条的所述时间信息进行修改，以缩短所述时间信息的表现形式；

可以理解的是，当日志是日志文件的第一条记录时，此时以该日志的时间信息为准，保留完整时间字符串，计算下一条日志的时间信息，即下一条日志的时间字符串与该日志的时间字符串的时间差值，并依据该时间差值修改此下一条日志的时间字符串，以缩短所述时间信息的表现形式；持续此环节则可节约大量的存储空间。

在另一些应用中，所述定义的日志内容标记优化规则，具体包括建立对应所述日志内容标记的映射关系值；基于所述映射关系值，对所述日志内容标记进行优化；

可以理解的是，在此应用中，建立对应日志内容标记的映射关系值，根据映射关系值即可替换相应的日志内容标记。

进一步的，基于定义的日志内容标记优化规则，对所述日志内容标记信息进行优化，以缩短所述内容标记信息的表现形式，具体包括建立对应所述日志内容标记的映射关系值；在所述日志数据库中依次保存与所述日志内容标记对应的映射关系值；依据所述各映射关系值，修改与所述各映射关系值对应的各所述日志内容标记，以缩短所述日志内容标记的表现形式；

可以理解的是，建立对应日志内容标记的映射关系值，将各映射关系值依次存在日志数据库中，通过查找对应日志内容标记进行替换，从而缩短日志内容标记的表现形式。

进一步的，依据所述各映射关系值，修改与所述各映射关系值对应的各所述日志内容标记，以缩短所述日志内容标记的表现形式后，所述方法还包括当接收到新日志时，判断是否存在与所述新日志的日志内容标记对应的所述映射关系值；当存在所述映射关系值时，修改所述新日志的日志内容标记，以缩短所述日志内容标记的表现形式；当不存在所述映射关系值时，建立与所述新日志的日志内容标记对应的所述映射关系值，并依据建立的所述映射关系值修改所述新日志的日志内容标记，以缩短所述日志内容标记的表现形式。

可以理解的是，为了避免新进入的日志占用过多的存储空间，当接收到新日志时，需要判断是否存在与新日志的日志内容标记对应的映射关系值，当存在时进行相应的修改，当不存在时，进行对应的创建，通过上述优化步骤，可最大化的节约存储空间。

S104，将所述优化后的时间信息以及日志内容标记信息在所述日志文件中进行更新存储。

可以理解的是，通过上述步骤可提高日志存储空间利用率，同样空间保存更多有效日志，保存日志时长也相应的增加，有利于后续的系统开发调试以及后续的车辆运营。

下面将结合图2和图3说明本发明日志文件存储方法在一些应用中的实施例：

为了便于说明问题，参考以下日志内容：

20 (2023-07-04 05:07:44.877852 529 2829 I ActivityManager: calcYSize:
 unsupported or RGB color format, RGBA8888_UBWC(-1040187392)
 2023-07-04 05:07:44.879030 3112 4878 I ComAdapterService:
 queryBinder + code is 32 binder is h.a.a.e.n@15334
 2023-07-04 05:07:44.880555 1020 2607 D CompatibilityChangeReporter:
25 Compat change id reported: 136219221; UID 1001000; state: DISABLED

2023-07-04 05:09:11.979622 5788 10328 I VehicleStateBinder:
mAdapterManager.getLowBeamHeight
2023-07-04 05:09:11.979694 5788 10328 I Proxy_CarVehicleStateManager:
getLowBeamHeight
5 2023-07-04 05:09:11.980545 3112 8876 I
VehicleSDK_CarVehicleStateManagerProxy: getLowBeamHeight

上述日志为标准 Android 日志格式，每行是一条记录，每条记录用空格分隔的多个字段，从左到右依次为：时间、进程号、线程号、日志等级、日志内容（日志内容开头冒号之前是日志标记[也叫日志 tag]）。

10 第一个字段是时间，例如：2023-07-04 05:07:44.877852，含义：2023 年 7 月 4 日,5 时 7 分 44 秒 877852 微秒。相邻 2 个日志记录之间的实际间隔很短，一般都在毫秒、甚至微秒级别。这个字段长度固定，为 26 个字节。

日志 tag，是各个日志输出方根据需要自己设置的名称，比如系统服务、应用等，以便区分不同模块输出的日志内容。可以看到这个字段长度大小不等。

15 在上述日志中，可以直观的看到时间字段和日志 tag 字段，在每行日志中占了不小的比重。

为了更好的理解实际情况，对实际的系统日志进行了统计分析，下表是统计结果：

日志记录	总行数	Tag 平均长度	Tag 字段所占 比例	时 间标记 长度	时间 标记所占 比例	Tag和 时间 总比 例
1	662692	12.9	0.115	26	0.231	0.346
2	310719	12.8	0.092	26	0.186	0.278

	5					
3	110749	10.8	0.088	26	0.213	0.301
	8					
平 均值	-	12.2	0.098	26	0.210	0.308

说明：

日志内容是日常真实开发日志，总共超过 1000 万行。在一定程度上，具有普遍代表性。

提取每行日志的 tag，计算长度和平均值，结果为 12.2，即平均每行中 tag 占 12 个字符（字节）以上。在所有日志文件中 tag 总共占比均值（所有 tag 总字节数÷日志总大小）为 0.098，约等于 10%。如果有 1G 的日志文件，其中 tag 约占用了 100M。

时间标记是固定长度（26 个字节），在日志文件中占比均值（时间字符串总字节数÷日志总大小）为 0.21，即超过 20%空间被时间字符串占用，如果有 1G 日志文件，超过 200M 空间被时间字符串占用。

Tag 和时间字符串总共占用的空间比例（二者求和）均值为 0.308，在 1G 日志中有近 1/3 的空间是被二者占用的。

下面说明优化方案：

如图 2 所示，对时间字符串的优化：

时间字符串的特点是每行时间都不同（有少量重复的），相邻时间差值很小。

每条日志记录不保存完整的时间字符串，只保存当前记录上和条记录的时间差，使用一个整数保存时间差。可根据需要（不同系统日志保存的频次）选择不同位数的整数类型，例如 2 字节、4 字节、8 字节等。

每个日志文件，第一行记录需要保存完整时间记录。后续日志时间，可在此基础上计算获得。

每个文件都会保存自身日志时间（第一行），保证每个日志文件可以独立使用，及时其他日志文件删除、丢失、损坏也不会收到影响。

5 如图 3 所示，对日志 tag 的优化：

日志 tag 的特点是：tag 总数相对有限（一个系统中通常只有几千个 tag），一个 tag 可能会多次重复出现。

保存一个 tag 映射文件，文件中保存：tag 名称，tag id（一个整数 id），最近更新时间。

10 日志记录过程中，把 tag 映射文件读入内容，可进行高效的读写。通过哈希算法检索 tag 并获取 tag id。哈希算法的时间复杂度为 $O(1)$ ，对运行效率几乎没有影响。

日志文件中不保存 tag 名称，只保存 tag id。每条新增日志从 tag 映射文件中检索获得对应的 tag id。一个 tag id 占 2 个字节，一个 tag 名称平均长度超过 15 12 个字节，这样用 2 个字节的 id 替换掉 12 个字节长的名称进行保存。

在还原日志内容时，通过日志映射文件查询到 id 对应的 tag 名称。

当一个 tag 超过一定的时长（可以设置固定时长，或者根据当前保存日志的最早时间），删除未使用的 tag，保证 tag 映射文件不会占用过大空间，仅保存对当前日志有效的 tag 信息。

20 优化效果分析：

时间字符串优化：

每个日志文件只有第一条记录保存完整时间。日志文件总数比较少，一般不超过 100 个，和成百上千万行的日志总量相比，每个文件第一行时间字符串所占空间可忽略。时间字符串最终优化为一个整数（时间差值），保存临近记

录之间的差值。即用 2 个字节（时间差）替代了 26 个字节（完整的时间字符串）。

日志 Tag 优化：

一个系统中的 tag 总数通常在几千个（或同等数量级），和成百上千万日志总量比，tag 映射文件的大小可忽略。日志 tag 最终优化为一个整数 id，通过这个 id 可以在 tag 映射文件中检索到对应的 tag 名称，用于还原日志内容。这样就用 2 个字节（tag id）替换了 12 个字节（tag 名称）。

性能和便利性

未使用复杂的压缩算法，在实时处理日志时，不会对系统性能带来负担。

10 由于算法的精细处理，几乎没有带来性能负担。

提高存储效率的同时，日志模块可以方便的处理和输出日志内容，便于实时查看。

下表总结了日志存储空间优化的效果：

字段	原长度	新长度	优化比例
时间	26字节	2字节	92%
日志Tag	12.2字节	2字节	84%

注：优化比例 = (原长度 - 新长度) ÷ 原长度

15 通过优化算法，时间字段优化节省了 92%的空间，在上述例子中，1G 日志文件中有超过 200M 被时间字段占用，可优化节省出 184M 的空间。日志 Tag 字段优化节省 84%的空间，在上述例子中有约 100M 空间被日志 Tag 占用，可优化节省约 84M 空间。整体上，1G 日志文件，通过优化总共可节省空间月 268M 空间，效果明显。

20 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 4 所示，本发明还提供一种日志文件存储系统，包括：

日志文件获取模块 201，配置为响应于日志存储优化请求，实时获取日志
5 数据库中包含的日志文件；

日志文件提取模块 202，配置为提取所述日志文件包含的时间信息以及日志内容标记信息；

日志文件优化模块 203，配置为基于定义的时间优化规则以及日志内容标记优化规则，对所述时间信息以及所述日志内容标记信息进行优化，以缩短所
10 述时间信息以及所述内容标记信息的表现形式；

日志文件存储模块 204，配置为将所述优化后的时间信息以及日志内容标记信息在所述日志文件中进行更新存储。

具体的，本实施例提供的日志文件存储系统，包括日志文件获取模块 201、日志文件提取模块 202、日志文件优化模块 203 以及日志文件存储模块 204。

15 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
20 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 5 所示，本发明还提供一种电子设备，包括：处理器、通信接口、存
25 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文件存储方法的步骤。

图5是本发明实施例提供的一种电子设备的结构示意图。如图5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；
5 该电子设备中的处理器710可以是一个或多个，图5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文件存储方法。

该电子设备还可以包括：输入装置730和输出装置740。

10 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5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文件存储方法对应的程序指令/模块。处理器710
15 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文件存储方法。

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
20 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。

5 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文件存储方法的步骤。

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、
10 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
15 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
20 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、一种日志文件存储方法，其特征在于，包括：

响应于日志存储优化请求，实时获取日志数据库中包含的日志文件；

提取所述日志文件包含的时间信息以及日志内容标记信息；

基于定义的时间优化规则以及日志内容标记优化规则，对所述时间信息以

5 及所述日志内容标记信息进行优化，以缩短所述时间信息以及所述内容标记信息的表现形式；

将所述优化后的时间信息以及日志内容标记信息在所述日志文件中进行更新存储。

2、根据权利要求 1 所述的日志文件存储方法，其特征在于，所述定义的时间优化规则，具体包括：

基于所述时间信息之间的时间差值，对所述时间信息进行优化。

3、根据权利要求 1 所述的日志文件存储方法，其特征在于，所述定义的日志内容标记优化规则，具体包括：

建立对应所述日志内容标记的映射关系值；

15 基于所述映射关系值，对所述日志内容标记进行优化。

4、根据权利要求 2 所述的日志文件存储方法，其特征在于，基于定义的时间优化规则，对所述时间信息进行优化，以缩短所述时间信息的表现形式，具体包括：

判断当前日志是否为所述日志文件的第一条记录；

20 当所述日志不是所述日志文件的第一条记录时，计算当前所述日志的所述时间信息与上一条日志的所述时间信息的差值；

依据所述时间信息的差值，对所述当前日志的所述时间信息进行修改，以缩短所述时间信息的表现形式。

5、根据权利要求4所述的日志文件存储方法，其特征在于，当所述日志是所述日志文件的第一条记录时，不对所述日志的时间信息进行优化，保留完整的所述时间信息；

计算所述日志的所述时间信息与下一条日志的所述时间信息的差值；

5 依据所述时间信息的差值，对所述下一条的所述时间信息进行修改，以缩短所述时间信息的表现形式。

6、根据权利要求3所述的日志文件存储方法，其特征在于，基于定义的日志内容标记优化规则，对所述日志内容标记信息进行优化，以缩短所述内容标记信息的表现形式，具体包括：

10 建立对应所述日志内容标记的映射关系值；

在所述日志数据库中依次保存与所述日志内容标记对应的映射关系值；

依据所述各映射关系值，修改与所述各映射关系值对应的各所述日志内容标记，以缩短所述日志内容标记的表现形式。

7、根据权利要求6所述的日志文件存储方法，其特征在于，依据所述各映射关系值，修改与所述各映射关系值对应的各所述日志内容标记，以缩短所述日志内容标记的表现形式后，所述方法还包括：

当接收到新日志时，判断是否存在与所述新日志的日志内容标记对应的所述映射关系值；

20 当存在所述映射关系值时，修改所述新日志的日志内容标记，以缩短所述日志内容标记的表现形式；

当不存在所述映射关系值时，建立与所述新日志的日志内容标记对应的所述映射关系值，并依据建立的所述映射关系值修改所述新日志的日志内容标记，以缩短所述日志内容标记的表现形式。

8、一种日志文件存储系统，其特征在于，包括：

日志文件获取模块，配置为响应于日志存储优化请求，实时获取日志数据库中包含的日志文件；

日志文件提取模块，配置为提取所述日志文件包含的时间信息以及日志内容标记信息；

- 5 日志文件优化模块，配置为基于定义的时间优化规则以及日志内容标记优化规则，对所述时间信息以及所述日志内容标记信息进行优化，以缩短所述时间信息以及所述内容标记信息的表现形式；

日志文件存储模块，配置为将所述优化后的时间信息以及日志内容标记信息在所述日志文件中进行更新存储。

- 10 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。

- 15 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。

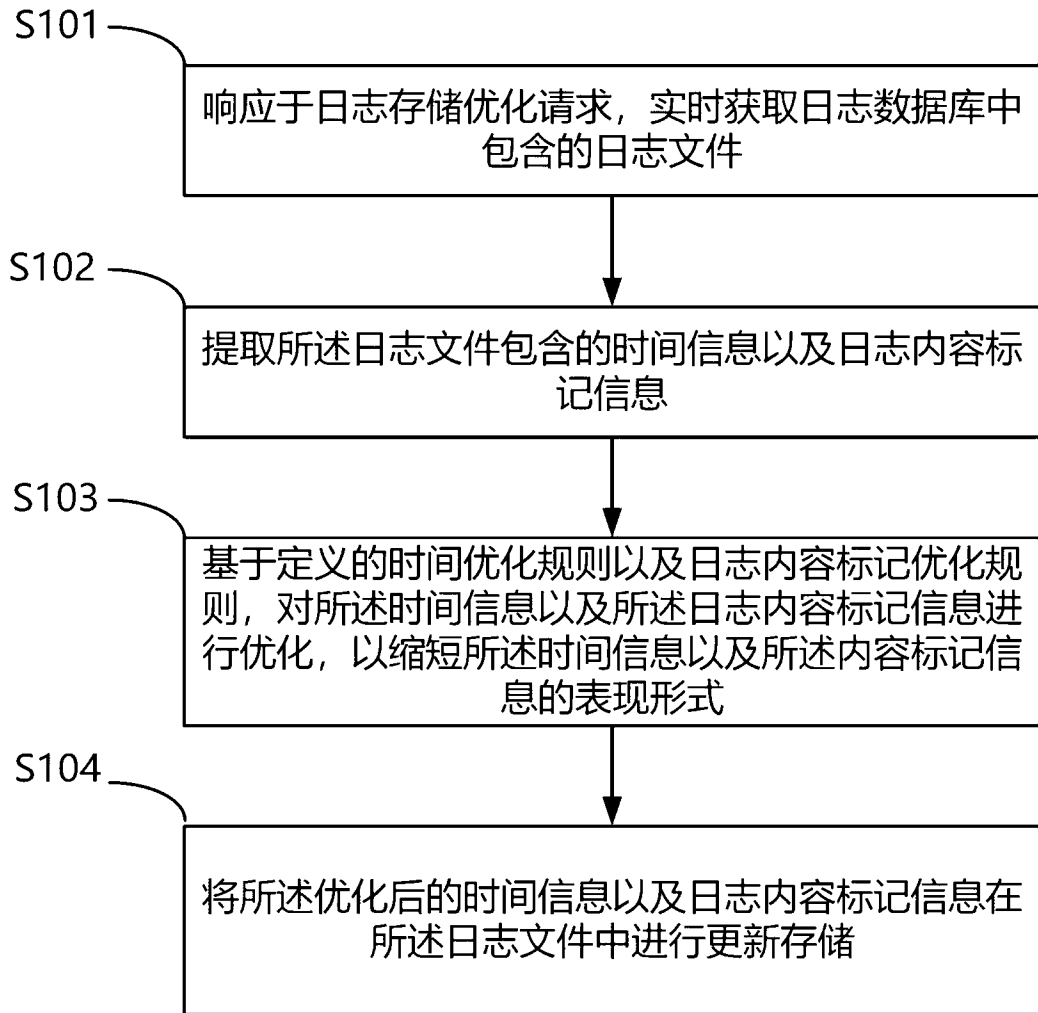


图 1

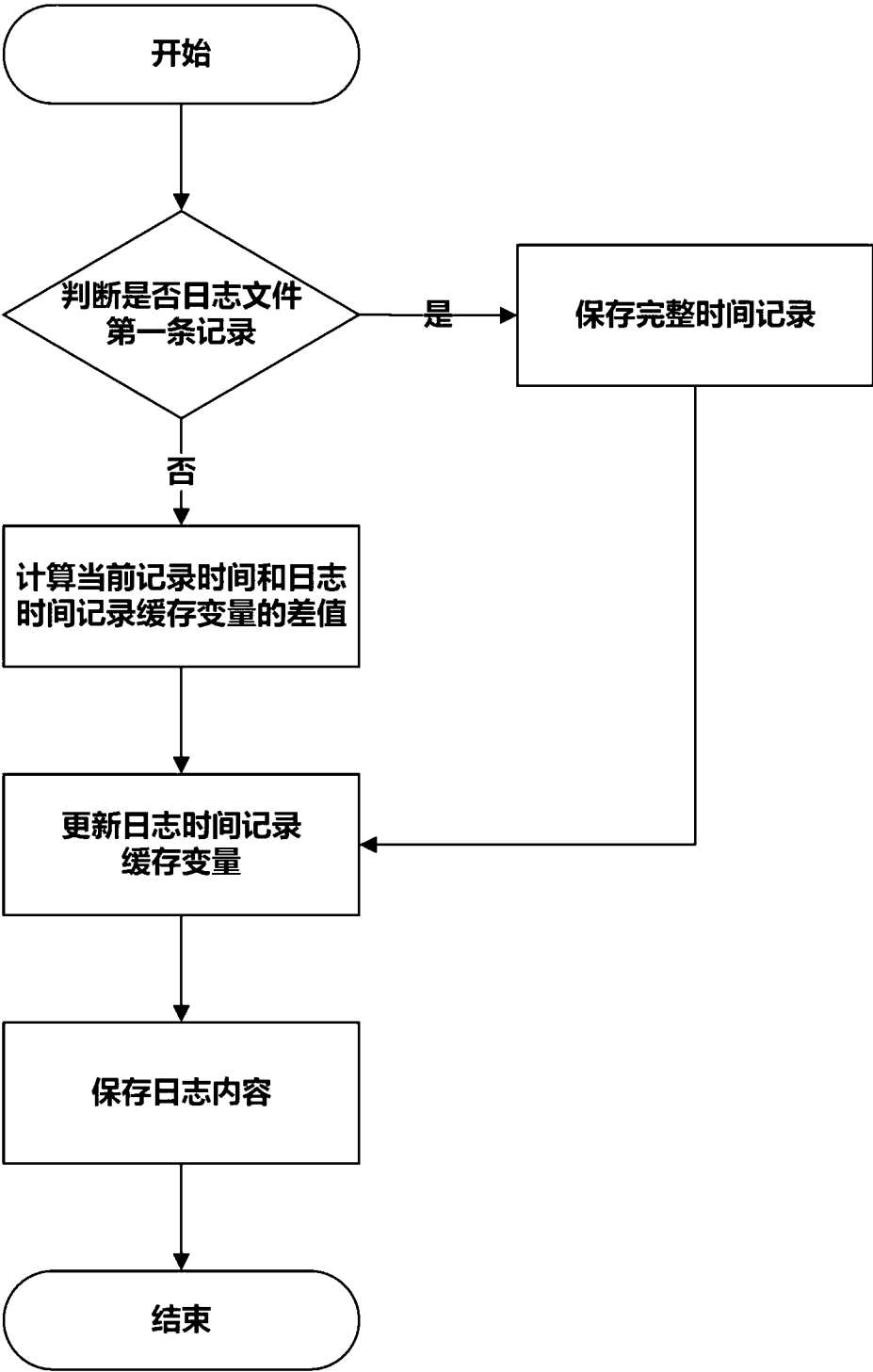


图 2

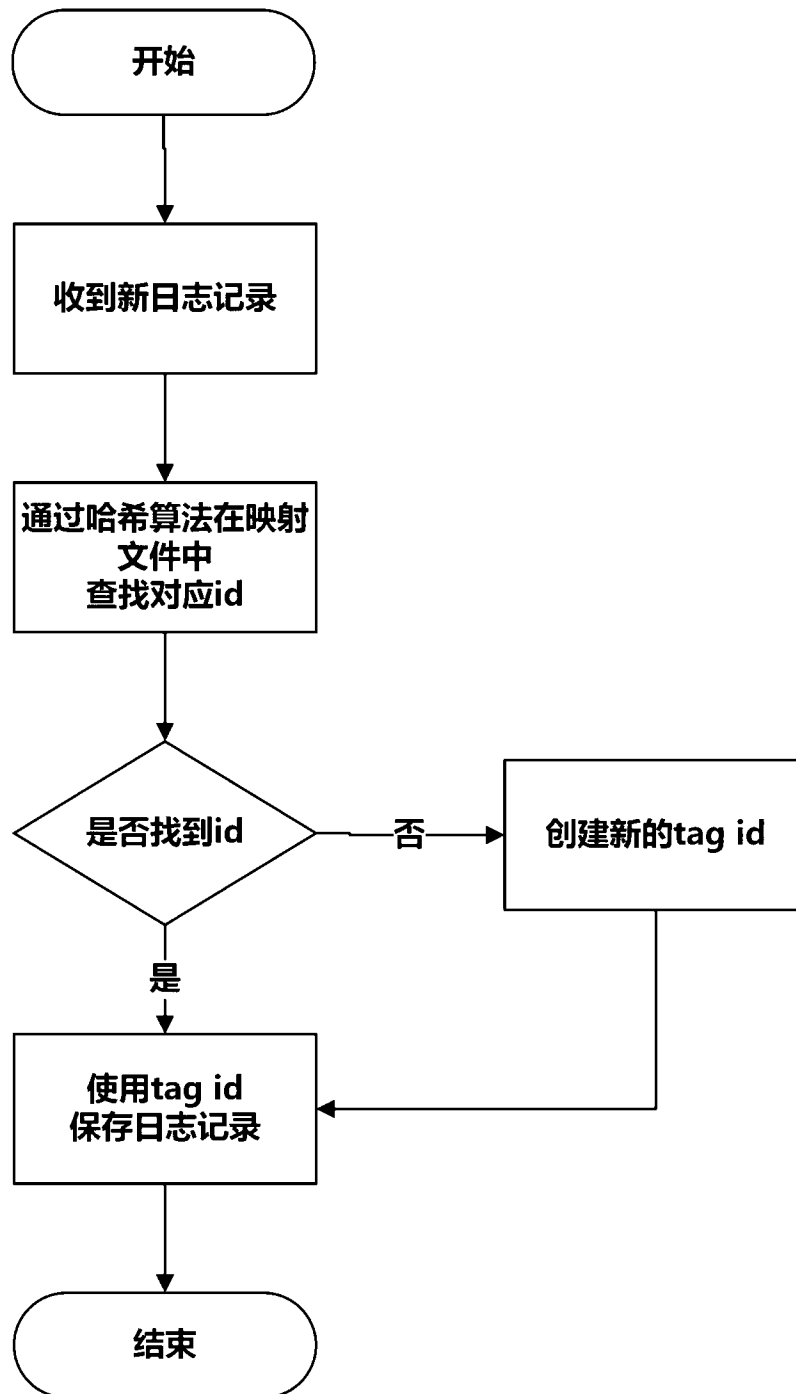


图 3

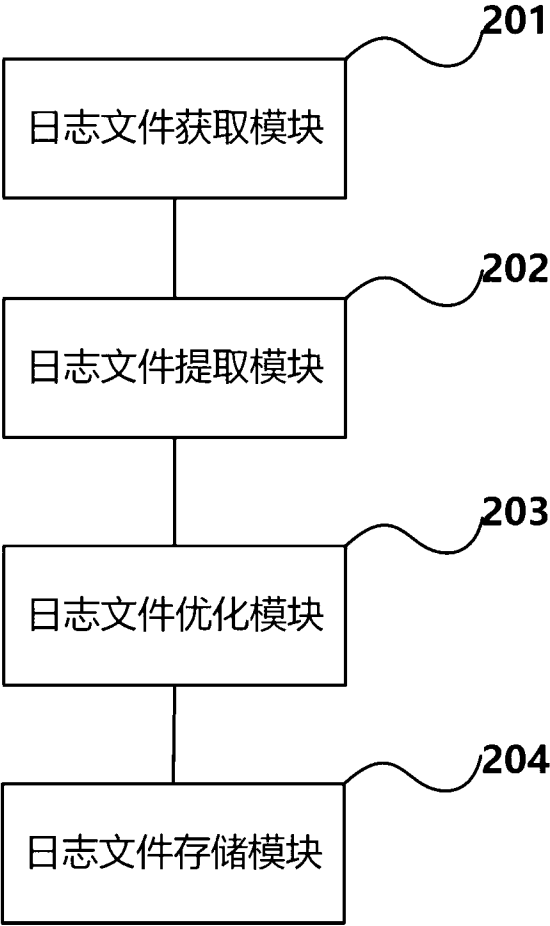


图 4

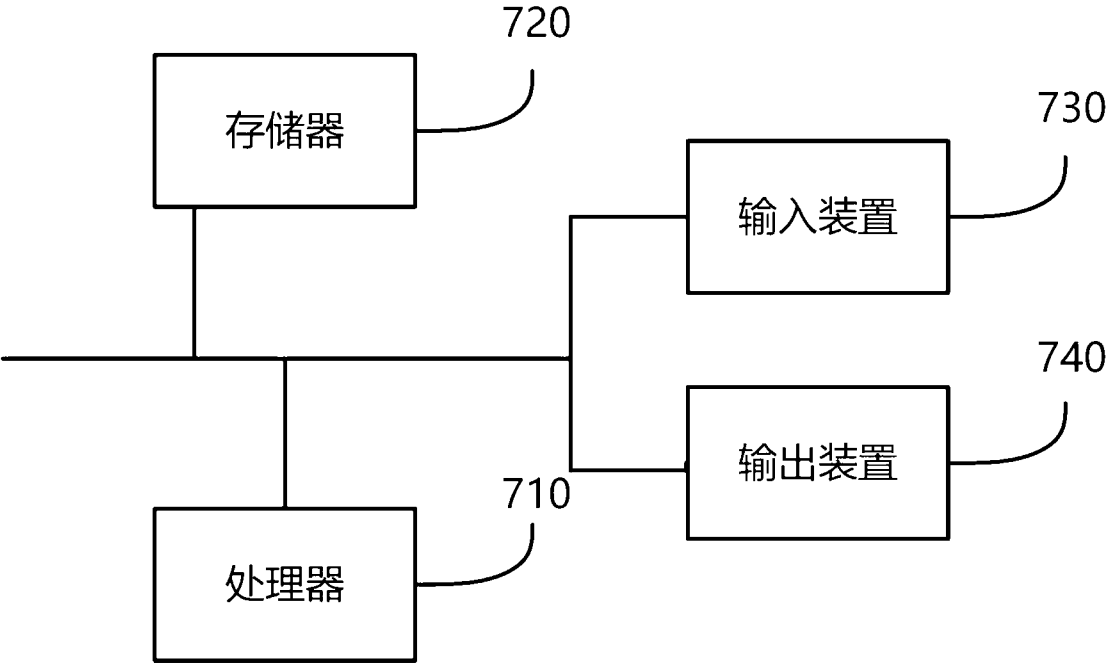


图 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2024/112984

A. CLASSIFICATION OF SUBJECT MATTER

G06F 16/18(2019.01)i; G06F 16/17(2019.01)i; G06F 16/13(2019.01)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)

CNABS, CNTXT, WOTXT, EPTXT, USTXT, CNKI, IEEE, VEN: 日志, 文件, 存储, 优化, 实时, 获取, 数据库, 时间, 内容标记, 规则, 缩减, 缩短, 压缩, 差值, 映射, 更新, log, file, store, optimize, real-time, obtain, database, time, content, tag, rule, reduce, shorten, compress, difference, map, update

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 117909309 A (CHINA FAW GROUP CORP. et al.) 19 April 2024 (2024-04-19) claims 1-10	1-10
X	CN 114860678 A (ZHEJIANG UNIVERSITY OF TECHNOLOGY) 05 August 2022 (2022-08-05) description, paragraphs [0032]-[0080], and figures 1-2	1-10
A	CN 111881012 A (SHENZHEN DASHI INTELLIGENT TECHNOLOGY CO., LTD.) 03 November 2020 (2020-11-03) entire document	1-10
A	US 2013111451 A1 (INTERNATIONAL BUSINESS MACHINES CORP.) 02 May 2013 (2013-05-02) 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

06 November 2024

Date of mailing of the international search report

13 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
Information on patent family members

International application No.
PCT/CN2024/112984

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	117909309	A	19 April 2024	None			
CN	114860678	A	05 August 2022	None			
CN	111881012	A	03 November 2020	None			
US	2013111451	A1	02 May 2013	US	8949799	B2	03 February 2015

A. 主题的分类

G06F 16/18(2019.01)i; G06F 16/17(2019.01)i; G06F 16/13(2019.01)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))

CNABS, CNTXT, WOTXT, EPTXT, USTXT, CNKI, IEEE, VEN: 日志, 文件, 存储, 优化, 实时, 获取, 数据库, 时间, 内容标记, 规则, 缩减, 缩短, 压缩, 差值, 映射, 更新, log, file, store, optimize, real-time, obtain, database, time, content, tag, rule, reduce, shorten, compress, difference, map, update

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 117909309 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
X	CN 114860678 A (浙江工业大学) 2022年8月5日 (2022 - 08 - 05) 说明书第[0032]-[0080]段, 附图1-2	1-10
A	CN 111881012 A (深圳大势智能科技有限公司) 2020年11月3日 (2020 - 11 - 03) 全文	1-10
A	US 2013111451 A1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 2013年5月2日 (2013 - 05 - 02) 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

2024年11月6日

国际检索报告邮寄日期

2024年11月13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

授权官员

徐淑娴

电话号码 (+86) 010-53961535

PCT/ISA/210 表(第2页) (2022年7月)

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2024/112984

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	117909309	A	2024年4月19日	无	
CN	114860678	A	2022年8月5日	无	
CN	111881012	A	2020年11月3日	无	
US	2013111451	A1	2013年5月2日	US 8949799 B2	2015年2月3日