

(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
2022 年 5 月 5 日 (05.05.2022)



(10) 国际公布号
WO 2022/089202 A1

(51) 国际专利分类号:
G06K 9/62 (2022.01)

中国广东省深圳市前海深港合作区前湾一路1号A栋201室, Guangdong 518052 (CN)。

(21) 国际申请号: PCT/CN2021/123363

(22) 国际申请日: 2021 年 10 月 12 日 (12.10.2021)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202011164795.1 2020年10月27日 (27.10.2020) CN

(71) 申请人: 深圳前海微众银行股份有限公司 (WEBANK CO., LTD.) [CN/CN]; 中国广东省深圳市前海深港合作区前湾一路1号A栋201室, Guangdong 518052 (CN)。

(72) 发明人: 杜鹤然 (DU, Heran); 中国广东省深圳市前海深港合作区前湾一路1号A栋201室, Guangdong 518052 (CN)。 胡盼盼 (HU, Panpan); 中国广东省深圳市前海深港合作区前湾一路1号A栋201室, Guangdong 518052 (CN)。 卢道和 (LU, Daohe);

(74) 代理人: 北京派特恩知识产权代理有限公司 (CHINA PAT INTELLECTUAL PROPERTY OFFICE); 中国北京市海淀区海淀南路21号中关村知识产权大厦B座2层, Beijing 100080 (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, IT, 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, 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, GH, GM, KE, LR, LS, MW, MZ,

(54) **Title:** FAULT IDENTIFICATION MODEL TRAINING METHOD, FAULT IDENTIFICATION METHOD, APPARATUS AND ELECTRONIC DEVICE

(54) 发明名称: 故障识别模型训练方法、故障识别方法、装置及电子设备

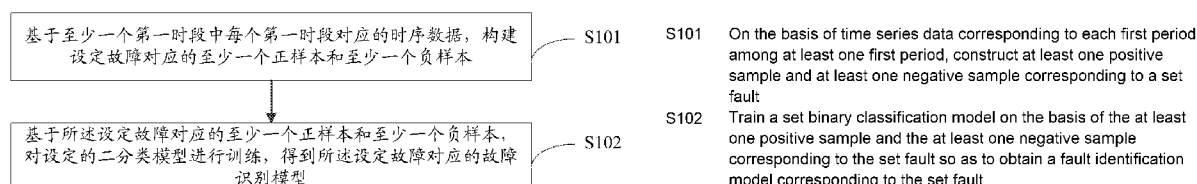


图 1

(57) **Abstract:** The embodiment of the present application discloses a fault identification model training method, a fault identification method, an apparatus and an electronic device. The fault identification model training method comprises: on the basis of time series data corresponding to each first period among at least one first period, constructing at least one positive sample and at least one negative sample corresponding to a set fault; training a set binary classification model on the basis of the at least one positive sample and the at least one negative sample corresponding to the set fault so as to obtain a fault identification model corresponding to the set fault, wherein the time series data comprises first time series data and second time series data; the first time series data represents the amount of log information outputted at each moment of the corresponding first period; and the second time series data represents monitoring data of each performance index in at least one performance index at each moment of the corresponding first period.

(57) **摘要:** 本申请实施例公开了一种故障识别模型训练方法、故障识别方法、装置及电子设备。该故障识别模型训练方法包括: 基于至少一个第一时段中每个第一时段对应的时序数据, 构建设定故障对应的至少一个正样本和至少一个负样本; 基于所述设定故障对应的至少一个正样本和至少一个负样本, 对设定的二分类模型进行训练, 得到所述设定故障对应的故障识别模型; 其中, 所述时序数据包括第一时序数据和第二时序数据; 第一时序数据表征对应的第一时段的每个时刻输出的日志信息量; 第二时序数据表征至少一个性能指标中每个性能指标在对应的第一时段的每个时刻的监测数据。

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))。

故障识别模型训练方法、故障识别方法、装置及电子设备

相关申请的交叉引用

本申请基于申请号为 202011164795.1，申请日为 2020 年 10 月 27 日的中国专利申请提出，并要求上述中国专利申请的优先权，上述中国专利申请的全部内容在此引入本申请作为参考。

技术领域

本申请涉及计算机技术领域，涉及一种故障识别模型训练方法、故障识别方法、装置及电子设备。

背景技术

随着计算机技术的发展，越来越多的技术应用在金融领域，传统金融业正在逐步向金融科技（Fintech）转变，然而，由于金融行业的安全性、实时性要求，金融科技也对技术提出了更高的要求。金融科技领域下，数据库即服务（DBaaS，Database as a Service）是以传统数据库技术为基础，将数据库资源以标准服务的形式提供给一个或多个租户的服务解决方案。

相关技术中，在用于提供 DbaaS 的服务器发生故障时，通常基于设定的性能指标的监测数据和 DbaaS 的相关日志对应的语义分析结果进行故障分析，得到分析结果。这种故障分析方法需要消耗大量时间分析所有日志文本的具体内容，以得到对应的语义分析结果，导致故障分析的效率较低。

发明内容

为解决相关技术问题，本申请实施例期望提供一种故障识别模型训练方法、故障识别方法、装置及电子设备。

本申请实施例提供一种训练故障识别模型的方法，包括：

基于至少一个第一时段中每个第一时段对应的时序数据，构建设定故障对应的至少一个正样本和至少一个负样本；

基于所述设定故障对应的至少一个正样本和至少一个负样本，对设定的二分类模型进行训练，得到所述设定故障对应的故障识别模型；其中，

所述时序数据包括第一时序数据和第二时序数据；

第一时序数据表征对应的第一时段的每个时刻输出的日志信息量；

第二时序数据表征至少一个性能指标中每个性能指标在对应的第一时段的每个时刻的监测数据。

本申请实施例还提供了一种故障识别方法，包括：

在检测到第一性能指标的数据触发告警的情况下，确定出第一时段对应的时序数据；

基于至少一种设定故障中每种设定故障对应的相关性能指标，确定出至少一种第一设定故障；所述第一设定故障的相关性能指标中包含所述第一性能指标；

5 将所述第一时段对应的时序数据输入所述至少一种第一设定故障中每种第一设定故障对应的故障识别模型，得到每个故障识别模型输出的识别结果；

将所有识别结果中置信度最高的第一设定故障确定为当前发生的故障；其中，所述第一时段的结束时刻为所述第一性能指标对应的告警时刻；

10 所述第一时段对应的时序数据包括实时日志对应的第一时序数据和至少一个性能指标中每个性能指标对应的第二时序数据；

所述故障识别模型基于上述任一种训练故障识别模型的方法训练得到。

上述方案中，所述故障识别方法还包括：

15 基于至少一台服务器中每台服务器发生的历史故障和每种历史故障对应的发生次数，确定出至少一台服务器中所有服务器的健康评分，以在健康评分大于或等于设定阈值的服务器中部署新的数据库实例。

本申请实施例还提供了一种训练故障识别模型的装置，包括：

构建单元，用于基于至少一个第一时段中每个第一时段对应的时序数据，构建设定故障对应的至少一个正样本和至少一个负样本；

20 训练单元，用于基于所述设定故障对应的至少一个正样本和至少一个负样本，对设定的二分类模型进行训练，得到所述设定故障对应的故障识别模型；其中，

所述时序数据包括第一时序数据和第二时序数据；

第一时序数据表征对应的第一时段的每个时刻输出的日志信息量；

第二时序数据表征至少一个性能指标中每个性能指标在对应的第一时段的每个时刻的监测数据。

25 本申请实施例还提供了一种故障识别装置，包括：

第一确定单元，配置为在检测到第一性能指标的数据触发告警的情况下，确定出第一时段对应的时序数据；

30 第二确定单元，配置为基于至少一种设定故障中每种设定故障对应的相关性能指标，确定出至少一种第一设定故障；所述第一设定故障的相关性能指标中包含所述第一性能指标；

第一识别单元，配置为将所述第一时段对应的时序数据输入所述至少一种第一设定故障中每种第一设定故障对应的故障识别模型，得到每个故障识别模型输出的识别结果；

35 第二识别单元，配置为将所有识别结果中置信度最高的第一设定故障确定为当前发生的故障；其中，

所述第一时段的结束时刻为所述第一性能指标对应的告警时刻；

所述第一时段对应的时序数据包括实时日志对应的第一时序数据和至少一个性能指标中每个性能指标对应的第二时序数据；

所述故障识别模型基于任一种训练故障识别模型的方法训练得到。

40 本申请实施例还提供了一种电子设备，包括：处理器和配置为存储能够在处理器上运行的计算机程序的存储器，

其中，所述处理器配置为运行所述计算机程序时，执行以下至少一项：

上述任一种训练故障识别模型的方法的步骤；

上述任一种故障识别方法的步骤。

本申请实施例还提供了一种存储介质，其上存储有计算机程序，所述计算机程序被处理器执行时以下至少一项：

上述任一种训练故障识别模型的方法的步骤；

上述任一种故障识别方法的步骤。

5 本申请实施例，基于至少一个第一时段中每个第一时段对应的时序数据，构建设定故障对应的至少一个正样本和至少一个负样本；基于所述设定故障对应的至少一个正样本和至少一个负样本，对设定的二分类模型进行训练，得到所述设定故障对应的故障识别模型。由于训练样本是基于第一时段对应的时序数据构建得到，第一时段对应的时序数据包括的第一时序数据表征第一时段的每个时刻输出的日志信息量，而日志信息量是指日志的信息量，用于度量日志传达的信息多少；在确定日志信息量时不需要分析历史日志的具体内容，也就是说，在构建训练样本时服务器不需要关注历史日志的具体内容，因此可以省去对日志的文本内容进行语义分析消耗的时间，提高获取训练样本的效率，进而提高故障识别模型的训练效率。

10 在利用故障识别模型进行故障识别时，由于输入故障识别模型的时序数据中包括实时日志对应的第一时序数据，第一时序数据表征每个时刻输出的日志信息量，在确定日志信息量时不需要对实时日志的文本内容进行语义分析，因此，可以节省时间，可以提高故障识别效率。

附图说明

图 1 为本申请实施例提供的一种训练故障识别模型的方法的实现流程示意图；

20 图 2 为本申请实施例提供的一种训练故障识别模型的方法中确定第一时序数据的实现流程示意图；

图 3 为本申请实施例提供的一种训练故障识别模型的方法中确定正样本和负样本的实现流程示意图；

25 图 4 为本申请实施例提供的一种训练故障识别模型的方法中确定正样本的实现流程示意图；

图 5 为本申请实施例提供的一种训练故障识别模型的方法中确定第一指标的实现流程示意图；

图 6 为本申请应用实施例提供的一种训练故障识别模型的方法的实现流程示意图；

图 7 为本申请实施例提供的一种故障识别方法的实现流程示意图；

30 图 8 为本申请另一实施例提供的一种故障识别方法的实现流程示意图；

图 9 为本申请实施例提供的训练故障识别模型的装置的结构示意图；

图 10 为本申请实施例提供的故障识别装置的结构示意图；

图 11 为本申请实施例提供的电子设备的硬件组成结构示意图。

具体实施方式

35 以下结合说明书附图及具体实施例对本申请的技术方案做进一步的详细阐述。

针对上述技术问题，本申请实施例提供了一种故障识别方法，基于日志信息量以及性能指标的监测数据进行故障识别。由于在确定日志信息量时不需要分析日志的具体内容，可以节省对日志的文本内容进行语义分析消耗的时间，提高故障识别效率。

以下结合说明书附图及具体实施例对本申请的技术方案做进一步的详细阐述。

40 图 1 示出了本申请实施例提供的一种训练故障识别模型的方法的实现流程示意图。

在本申请实施例中，训练故障识别模型的方法的执行主体为电子设备，例如，电脑、服务器等。

参照图 1，本申请实施例提供的训练故障识别模型的方法包括：

S101：基于至少一个第一时段中每个第一时段对应的时序数据，构建设定故障对应的至少一个正样本和至少一个负样本；其中，所述时序数据包括第一时序数据和第二时序数据；第一时序数据表征对应的第一时段的每个时刻输出的日志信息量；第二时序数据表征至少一个性能指标中每个性能指标在对应的第一时段的每个时刻的监测数据。

这里，第一时段包括发生设定故障的时段和未发生该设定故障的时段。

电子设备基于第一时段中发生设定故障的时段对应的第一时序数据和第二时序数据构建正样本；基于第一时段中未发生设定故障的时段对应的第一时序数据和第二时序数据构建负样本。

需要说明的是，发生设定故障时，至少一个性能指标的数据触发告警。电子设备可以基于在发生该设定故障时首个触发告警的性能指标对应的告警时刻，确定出发生该设定故障的时段以及确定出未发生该设定故障的时段。发生该设定故障的时段包括首个告警时刻；未发生该设定故障的时段为对应的第一时段中除发生该设定故障的时段之外的时段。

例如，设定故障对应的至少一个性能指标中的第一性能指标在第一时段中的第一时刻触发告警，电子设备可以将第一时刻之前的第二时刻确定为发生设定故障的时段的开始时刻，将第一时刻之后的第三时刻确定为发生设定故障的时段的结束时刻。需要说明的是，第二时刻在对应的第一时段的开始时刻之后，第三时刻在对应的第一时段的结束时刻之前。第一时长和第二时长可以相同，也可以不同。第一时长为第一时刻与第二时刻之间的时长，第二时长为第一时刻与第三时刻之间的时长。示例性地，第一时长和第二时长可以为 15 分钟，当然，也可以根据实际情况进行设置。

需要说明的是，第一时序数据是按时间先后顺序记录的日志信息量。第二时序数据是按时间先后顺序记录的至少一个性能指标中每个性能指标的监测数据。每个性能指标对应一组第二时序数据。其中，第一时序数据基于 DBaaS 服务器的历史日志得到。第二时序数据基于 DBaaS 服务器的性能指标的监测数据得到。至少一个性能指标用于监测是否发生设定故障。设定故障可以为连接数饱和、机器磁盘故障、内存故障、慢查询过多等。

日志信息量是指日志的信息量。日志信息量用于量度日志的信息多少。每个时刻输出的日志信息量是该时刻打印出的所有日志中每条日志的信息量的总和。一条日志的信息量基于第一概率和至少一个第二概率确定出。第一概率表征该条日志对应的日志级别出现的概率。第二概率表征该条日志中包括的所有设定词组中每个设定词组在该日志对应的日志级别下出现的概率。

示例性地，日志级别可以包括 FATAL、ERROR、WARN、INFO、DEBUG。其中，FATAL 级别的日志表征每个严重的错误事件将会导致应用程序退出。

ERROR 级别的日志表征虽然发生错误事件，但仍然不影响系统继续运行。

WARN 级别的日志表征系统存在潜在的错误事件。

INFO 级别的日志表征应用程序的运行情况。

DEBUG 级别的日志主要用于在调试应用程序时更详细的了解系统的运行状态。

S102：基于所述设定故障对应的至少一个正样本和至少一个负样本，对设定的二分类模型进行训练，得到所述设定故障对应的故障识别模型。

电子设备将至少一个正样本中的每个正样本转换成对应的第一向量，将至少一个负样本中的每个负样本转换成对应的第二向量；将至少一个第一向量和至少一个第二向量

输入至设定的二分类模型进行训练,得到所述设定故障对应的故障识别模型。

其中,在训练设定的二分类模型的过程中,在二分类模型未达到设定的收敛条件时,更新二分类模型的相关参数,基于至少一个正样本和至少一个负样本继续训练该二分类模型。设定的收敛条件可以是第一模型参数与第二模型参数之间的差值小于或等于设定
5 阈值,当然也可以设置其他收敛条件,例如,训练次数达到设定次数。第一模型参数表征第 k 次迭代训练对应的模型参数,第二模型参数表征第 $k-1$ 次迭代训练对应的模型参数, k 为大于或等于 1 的整数。

在二分类模型达到设定的收敛条件时,,停止训练,将最后一次更新得到的模型参数确定为训练完毕的二分类模型所使用的模型参数,并将训练完毕的二分类模型确定为
10 设定故障对应的故障识别模型。

需要说明的是,由于设定故障对应的正样本和负样本的数量通常比较少,为了防止过拟合,设定的二分类模型通常为简单的卷积神经网络,例如,由一层卷积层和隐藏层构成的卷积神经网络。设定的二分类模型也可以为决策树模型。

在本实施例提供的方案中,基于至少一个第一时段中每个第一时段对应的第一时序数据和第二时序数据,构建设定故障对应的至少一个正样本和至少一个负样本;基于所述设定故障对应的至少一个正样本和至少一个负样本,对设定的二分类模型进行训练,得到所述设定故障对应的故障识别模型。由于第一时序数据表征对应的第一时段的每个
15 时刻输出的日志信息量,而日志信息量是指日志的信息量,用于度量日志传达的信息多少;在确定日志信息量时不需要分析历史日志的具体内容,因此可以省去对日志的文本内容进行语义分析消耗的时间,提高获取训练样本的效率,进而提高故障识别模型的训练效率。

在一实施例中,图 2 示出了本申请实施例提供的一种训练故障识别模型的方法中确定第一时序数据的实现流程示意图。参照图 2,确定第一时序数据的方法包括:

S201: 基于第一时段中每个时刻打印出的每条日志对应的第一信息量和第二信息
25 量,计算出每条日志对应的总信息量;其中,所述第一信息量表征日志对应的日志级别出现的第一概率;所述第二信息量表征日志中包括的所有设定词组中每个设定词组在对应的日志级别中出现的第二概率。

电子设备计算每条日志对应的第一信息量和第二信息量之和,得到对应日志的总信息量。

30 这里,可以基于以下公式计算每条日志对应的总信息量:

$$H(X) = H(x_i) + H(x_c) \quad (1)$$

$$H(x_i) = -\log P(x_i) \quad (2)$$

$$H(x_c) = \sum_{i=1}^n -\log(P(x_i | x_i)) \quad (3)$$

其中, $H(X)$ 表征一条日志的总信息量, $H(x_i)$ 表征日志的第一信息量, $H(x_c)$ 表征
35 日志的第二信息量; $P(x_i)$ 表征日志对应的日志级别 x_i 出现的第一概率; $P(x_i | x_i)$ 表征在日志对应的日志级别 x_i 中第 i 个设定词组出现的第二概率, n 表征对应日志中包含的设定词组的数量。

在一实施例中,按照以下方式确定出第一概率和第二概率:

基于至少一个第一日志样本中每个日志级别出现的次数,确定出每个日志级别对应
40 的第一概率;

基于所述至少一个第一日志样本中每个设定词组在每个日志级别下出现的次数,确定出每个设定词组在对应的日志级别中对应的第二概率;其中,

所述至少一个第一日志样本通过对历史日志进行采样得到。

这里，电子设备对采集到的历史日志进行采样，得到至少一个第一日志样本。采集到的历史日志可以包括至少一个第一时段中每个第一时段输出的历史日志中的部分或全部，采集到的历史日志也可以不包括第一时段输出的历史日志。

5 在一实施例中，至少一个第一日志样本中每个第一日志样本均满足以下条件：

第一日志样本对应的日志级别对应的第一概率满足设定条件；

第一日志样本对应的告警类型为发生所述设定故障时监测到的告警类型；

10 第三概率与第四概率之间的差值小于或等于设定阈值；其中，所述第三概率表征第一日志样本对应的告警类型在所述至少一个第一日志样本中出现的概率；所述第四概率表征第一日志样本对应的告警类型在发生所述设定故障时被监测到的概率。

15 其中，设定条件表征为所述设定故障配置的所有日志级别中每种日志级别对应的概率范围。由于数据库中的设定组件的源码决定了在发生设定故障时输出哪种日志级别的日志，以及决定了每种日志级别出现的概率，因此，电子设备可以分析数据库中的设定组件的源码，确定出在设定故障时输出的日志对应的每种日志级别对应的概率范围，基于确定出的概率范围设置上述设定条件。

第一日志样本对应的所有告警类型由第一日志样本中 ERROR 级别的日志中的设定词组确定出；一条日志包含至少一个设定词组；一个设定词组对应一个告警类型。

电子设备在得到至少一个第一日志样本的情况下，基于至少一个第一日志样本中的所有第一日志样本确定出第一概率和第二概率，具体如下：

20 电子设备基于第一日志样本包括的至少一条日志中每条日志对应的日志级别，统计出所有第一日志样本中每种日志级别出现的次数；对所有第一日志样本中每种日志级别出现的次数进行求和运算，得到所有第一日志样本中所有日志级别出现的第一总次数；基于所有第一日志样本中每种日志级别出现的次数，以及基于所有第一日志样本中所有日志级别出现的第一总次数，计算出每种日志级别对应的第一概率。第一概率为对应的

25 日志级别出现的次数与第一总次数的商。

电子设备基于分词技术，对第一日志样本中的每条日志进行分词，得到分词结果，对分词结果进行过滤处理，提取出每条日志对应的设定词组。对分词结果进行过滤处理包括去除连接词、常用词（例如，end、run 等）、组件的名称、调用的脚本的名称等。设定词组表征设定故障对应的告警类型（或称告警事件）。例如，设定故障为连接数饱和时，设定词组可以包括：operation timeout、checkWillUpdateKafka error 等。

30

电子设备基于所有第一日志样本中每条日志对应的设定词组，统计出每个设定词组在对应的日志级别中出现的次数；对每个设定词组在对应的日志级别中出现的次数进行求和运算，统计出所有设定词组对应的日志级别中出现的第二总次数；基于每个设定词组在对应的日志级别中出现的次数，以及基于对应的第二总次数，确定出每个设定词组

35 在对应的日志级别中出现的第二概率。其中，一条日志对应一个设定的日志级别，第二概率为每个设定词组在对应的日志级别下出现的次数与对应的第二总次数的商。

在本实施例中，第一日志样本满足上述三个条件，第一日志样本能够准确反映出发生设定故障时服务器的运行情况，利用基于第一日志样本确定出的第一概率和第二概率计算出的日志信息量，能够准确反映出发生设定故障时的日志信息量，进而提高训练得到的故障识别模型的准确度。

40

S202：基于第一时段中每个时刻打印出的每条日志对应的总信息量，得到对应的第一时段对应的第一时序数据。

电子设备对同一时刻打印出的每条日志对应的总信息量进行求和运算，得到对应时刻打印出的所有日志对应的总信息量；基于第一时段中每个时刻打印出的所有日志对应

的总信息量，输出对应的第一时段对应的第一时序数据。其中，同一时刻打印出的所有日志对应的总信息量，即为对应时刻输出的日志信息量。

每个时刻打印出的所有日志对应的总信息量 $H(x_s)$ 为：

$$H(x_s) = \sum_{k=1}^m H(X_k) \quad (4)$$

其中， m 表征对应时刻打印出的所有日志的数量， $H(X_k)$ 表征对应时刻打印出的第 k 条日志的总信息量。

这里，将通过上述公式 (1) 至 (3) 计算出的每条日志的总信息量，代入公式 (4)，即可得到对应时刻打印出的所有日志对应的总信息量。

例如，在第一时段中第一时刻打印出了两条“ERROR”级别的日志，第一条日志包括一个设定词组“operation timeout”，第二条日志包括一个设定词组“checkWillUpdateKafka error”，预先基于至少一个第一日志样本确定出“ERROR”级别出现的概率为 1%，“operation timeout”在 ERROR 级别中出现的概率为 10%，“checkWillUpdateKafka error”出现的概率为 5%，那么基于上述公式 (4) 可以得到第一时刻输出的信息量的表达式为：

$$H(x_s) = \sum_{k=1}^2 H(X_k) = H(X_1) + H(X_2) \quad (5)$$

基于上述 (1) 至 (3) 可以得到：

$$H(X_1) = -\log P(0.01) - \log P(0.1) \quad (6)$$

$$H(X_2) = -\log P(0.01) - \log P(0.05) \quad (7)$$

将 (6) 和 (7) 代入 (5)，得到第一时刻输出的信息量为：

$$H(x_s) = -\log P(0.01) - \log P(0.1) - \log P(0.01) - \log P(0.05) \approx 6.3$$

在本实施例提供的方案中，基于第一时段中每个时刻打印出的每条日志对应的第一信息量和第二信息量，计算出每条日志对应的总信息量；基于第一时段中每个时刻打印出的每条日志对应的总信息量，得到对应的第一时段对应的第一时序数据。由于可以准确计算出每条日志对应的总信息量，因此提高第一时序数据的准确度，进而提高训练得到的故障识别模型的准确度。

作为本申请的另一实施例，图 3 示出了本申请实施例提供的一种训练故障识别模型的方法中确定正样本和负样本的实现流程示意图。参照图 3，所述基于至少一个第一时段对应的时序数据，构建设定故障对应的至少一个正样本和至少一个负样本，包括：

S301：基于第一时段中的第二时段对应的时序数据，确定出与所述设定故障相关的第一指标；其中，所述第二时段表征对应的第一时段中发生所述设定故障的时段。

这里，所述第一指标包括以下至少之一：

在发生所述设定故障时告警的所有性能指标中的至少一个性能指标；

日志信息量。

由于在发生设定故障时，存在至少一个性能指标的监测数据触发告警，因此，电子设备可以基于在发生设定故障时触发告警的性能指标对应的告警时刻，确定出对应的第一时段中的第二时段。其中，第二时段的起始时刻小于或等于在发生设定故障时的首个告警时刻；第二时段的结束时刻大于或等于在发生设定故障时的最后一个告警时刻。

例如，将第一个告警时刻以及该告警时刻前后 15 分钟对应的时段，确定为对应的第一时段中的第二时段。

S302：基于所述第一指标在所述第二时段对应的时序数据，构建所述设定故障对应的至少一个正样本。

这里,可基于第一指标在一个第二时段对应的时序数据,构建一个正样本。

在一实施例中,图4示出了本申请实施例提供的一种训练故障识别模型的方法中确定正样本的实现流程示意图。参照图4,所述基于所述第一指标在所述第二时段对应的时序数据,构建所述设定故障对应的至少一个正样本,包括:

- 5 S401: 基于所述第一指标的第一数量,确定出正样本的第二数量;其中,所述第一数量为大于或等于2的整数,所述第二数量为正整数;所述第二数量小于或等于所述第一数量对应的全排列运算结果的二分之一。

这里,通过公式 $M \leq \frac{A_N^N}{2}$ 确定出第二数量。

- 10 其中, M 表征正样本的数量, N 表征第一指标的数量。M 为正整数, N 大于或等于2的整数。

S402: 将所有第一指标在第二时段对应的时序数据以第一指标为最小单位进行全排列,得到全排列结果。

- 15 这里,电子设备将每个第一指标在第二时段对应的时序数据,按照时间先后顺序进行排列,得到对应的序列。将每个第一指标对应的序列以第一指标为最小单位进行全排列,得到全排列结果。

- 20 需要说明的是,通过变换所有第一指标的时序数据的排列顺序以扩充正样本,可以节省选取正样本消耗的时间。考虑到在发生设定故障时,对应的性能指标的告警顺序并不是固定不变的,因此,变换所有第一指标的时序数据的排列顺序,还可以使得训练得到的故障识别模型对第一指标对应的时序数据的排列顺序不敏感,提升故障识别模型的泛化能力。

S403: 从所述全排列结果中选出第二数量的正样本。

这里,一个正样本对应一种排列方式的序列。需要说明的是,每个正样本也可以为一维矩阵,一个一维矩阵对应一种排列方式的序列。

- 25 示例性地,第一指标的数量为3时,第一指标的全排列结果为6,基于3个第一指标最多可得到3个正样本。

S303: 基于所述第一指标在所述第一时段中的第三时段对应的时序数据,构建所述设定故障对应的至少一个负样本;其中,所述第三时段表征对应的第一时段中未发生所述设定故障的时段。

- 30 第三时段为对应的第一时段中除第二时段之外的任一时段。这里,可以基于第一指标在一个第三时段对应的时序数据,构建一个负样本。

- 35 本实施例提供的方案中,通过与设定故障相关的第一指标在第一时段对应的时序数据,构建至少一个正样本以及至少一个负样本,可以提高构建出的正样本与设定故障的相关性,进而提高基于正样本和负样本训练得到的故障识别模型的性能,进而提高训练得到的故障识别模型的准确度。在利用该故障识别模型进行故障识别时,可以提高识别结果的准确度。

作为本申请的另一实施例,图5示出了本申请实施例提供的一种训练故障识别模型的方法中确定第一指标的实现流程示意图。参照图5,所述基于第一时段中的第二时段对应的时序数据,确定出与所述设定故障相关的第一指标,包括:

- 40 S501: 基于第二时段对应的时序数据,确定出至少两个集合。

其中,所述至少两个集合包括一个第一集合和至少一个性能指标中每个性能指标对应的第二集合;所述第一集合中的元素表征相邻两个时刻输出的日志信息量之间的差值;所述第二集合中的元素表征对应的性能指标在相邻两个时刻的监测数据之间的差值。

这里，基于第二时段对应的第一时序数据，确定出第一集合；基于至少一个性能指标中每个性能指标在第二时段对应的第二时序数据，确定出对应的第二集合。至少两个集合中每个集合中的所有元素按照时间先后顺序排列。

S502：计算出所述至少两个集合中每个集合对应的均值和标准差。

5 电子设备可以基于均值计算公式，计算出至少两个集合中每个集合对应的均值；基于标准差的计算公式，计算出至少两个集合中每个集合对应标准差。

S503：基于三西格玛准则以及基于计算出的均值和标准差，确定出所述至少两个集合中每个集合对应的第一区间。

10 三西格玛准则指出，符合正态分布的一组数据中的数据分布在 $(\mu - 3\sigma, \mu + 3\sigma)$ 中的概率为 99.73%，分布在 $(\mu - 3\sigma, \mu + 3\sigma)$ 之外的概率为 0.27%。

因此，将 $(\mu - 3\sigma, \mu + 3\sigma)$ 确定为对应的第一区间。 μ 表征对应集合的均值； σ 表征对应集合的标准差。

15 S504：基于所述至少两个集合中每个集合对应的第一区间和对应集合中的最大值，确定出与所述设定故障相关的第一指标；其中，所述第一指标对应的集合中的最大值未处于对应的第一区间。

这里，当第一集合中的最大值未处于第一集合对应的第一区间时，将日志信息量确定为与设定故障相关的第一指标。当第一集合中的最大值处于第一集合对应的第一区间时，表征日志信息量不是与设定故障相关的第一指标。

20 当第二集合中的最大值未处于对应的第一区间时，将该第二集合对应的性能指标确定为设定故障相关的第一指标。当第二集合中的最大值处于对应的第一区间时，表征该第二集合对应的性能指标不是与设定故障相关的第一指标。

25 需要说明的是，当第一时段的数量为至少两个时，可以按照上述方式确定出至少两份第一指标，将至少两份第一指标中的交集确定为最终的第一指标，这样可以提高第一指标的准确度。其中，每份第一指标基于一个第一时段中的第二时段对应的时序数据确定出。

本实施例提供的方案中，基于三西格玛准则确定出对应的第一区间，基于确定出的第一区间筛选出与设定故障相关的第一指标，可以提高确定出的第一指标的准确度。

30 需要说明的是，图 5 对应的实施例中是基于第二时段对应的所有时序数据以及三西格玛准则确定出与设定故障相关的第一指标。在另一实施例中，所述第二时段对应的时序数据未包括在第二时段内首个触发告警的第一性能指标对应的第二时序数据；所述基于所述至少两个集合中每个集合对应的第一区间和对应集合中的最大值，确定出与所述设定故障相关的第一指标，包括：

基于所述至少两个集合中每个集合对应的第一区间和对应集合中的最大值，确定出第二指标；其中，所述第二指标对应的集合中的最大值未处于对应的第一区间；

35 将确定出的第二指标和所述第一性能指标确定为与所述设定故障相关的第一指标。

这里，电子设备基于第二时段对应的第二时序数据，确定出在对应的第二时段内首个触发告警的第一性能指标；基于第二时段对应的第一时序数据确定出第一集合；基于第二性能指标在第二时段对应的时序数据，确定出对应的第二性能指标对应的第二集合。其中，第二性能指标表征至少一个性能指标中除第一性能指标之外的任一性能指标。

40 当第一集合中的最大值未处于第一集合对应的第一区间时，将日志信息量确定为与设定故障相关的第二指标。当第一集合中的最大值处于第一集合对应的第一区间时，表征日志信息量不是与设定故障相关的第二指标。

当第二集合中的最大值未处于对应的第一区间时，将该第二集合对应的性能指标确定为与设定故障相关的第二指标。当第二集合中的最大值处于对应的第一区间时，表征

该第二集合对应的性能指标不是与设定故障相关的第二指标。

在确定出所有第二指标的情况下，电子设备将第一性能指标和确定出的所有第二指标，确定为与设定故障相关的第一指标。

本实施例提供的方案中，将在对应的第二时段内首个触发告警的第一性能指标识别为与设定故障相关的第一指标，然后再按照上述方法确定出与设定故障相关的第二指标，将第一性能指标和确定出的第二指标识别为与设定故障相关的第一指标。不需要通过第一性能指标对应的第二时序数据确定第一性能指标是否与设定故障相关，可以节省确定第一指标的时间，提高故障识别模型的训练效率。

作为本申请的应用实施例，图6示出了本申请应用实施例提供的一种训练故障识别模型的方法的实现流程示意图。参照图6，训练故障识别模型的方法包括：

S601：按照设定的时间间隔对历史日志进行采样，得到至少一个第一日志样本。

设定的时间间隔可以为1秒。

其中，第一日志样本中每个第一日志样本均满足以下条件：

第一日志样本对应的日志级别对应的第一概率满足设定条件；

第一日志样本对应的告警类型为发生所述设定故障时监测到的告警类型；

第三概率与第四概率之间的差值小于或等于设定阈值；其中，所述第三概率表征第一日志样本对应的告警类型在所述至少一个第一日志样本中出现的概率；所述第四概率表征第一日志样本对应的告警类型在发生所述设定故障时被监测到的概率。

S602：基于所述至少一个第一日志样本，确定出至少一个第一概率和至少一个第二概率；其中，第一概率为每种日志级别对应的概率；第二概率为每个设定词组在对应的日志级别中对应的第二概率。

S602的实现过程请参照上述实施例中S201中的相关描述，此处不赘述。

S603：基于确定出的第一概率和第二概率，确定出历史日志对应的时序数据。

这里，电子设备基于确定出的第一概率和第二概率，确定出每个时刻打印出的历史日志中每条日志的总信息量；基于历史日志中每条日志的总信息量，确定出每个时刻输出的日志信息量，从而得到历史日志对应的时序数据。确定历史日志对应的时序数据实现方式请参照上述实施例S201~S202中的相关描述，此处不赘述。

S604：基于至少一个第一时段中每个第一时段对应的时序数据，确定出与设定故障相关的第一指标。

这里，每个第一时段对应的时序数据包括第一时序数据和第二时序数据。第一时序数据表征对应的第一时段的每个时刻输出的日志信息量；第二时序数据表征至少一个性能指标中每个性能指标在对应的第一时段的每个时刻的监测数据。第二时序数据基于监测程序监测到的至少一个性能指标中每个性能指标在每个时刻对应的监测数据得到。

S604的实现方式请参照上述S301中的相关描述，此处不赘述。

S605：基于所述第一指标在第一时段中的第二时段对应的时序数据，构建设定故障对应的至少一个正样本。

这里，S605的实现方式请参照上述S302中的相关描述，此处不赘述。

S606：基于所述第一指标在第一时段中的第三时段对应的时序数据，构建设定故障对应的至少一个负样本。

这里，S606的实现方式请参照上述S303中的相关描述，此处不赘述。

S607：基于所述设定故障对应的至少一个正样本和至少一个负样本，对设定的二分类模型进行训练，得到所述设定故障对应的故障识别模型。

这里，S607的实现方式请参照上述S102中的相关描述，此处不赘述。

本实施例提供的方案中，基于所有第一日志样本确定出至少一个第一概率和至少一

个第二概率；基于确定出的第一概率和第二概率，确定出至少一个第一时段中每个第一时段对应的第一时序数据；基于第一时段对应的第一时序数据和第二时序数据，确定出设定故障对应的至少一个正样本和至少一个负样本；基于至少一个正样本和至少一个负样本训练得到设定故障对应的故障识别模型。由于第一日志样本能够准确反映出发生设定故障时服务器的运行情况，因此，利用基于第一日志样本确定出的概率计算出的日志信息量，能够准确反映出发生设定故障时的日志信息量，进而提高训练得到的故障识别模型的准确度。由于第一时序数据表征第一时段的每个时刻输出的日志信息量，而日志信息量是指日志的信息量，用于度量日志传达的信息多少；在确定日志信息量时不需要分析历史日志的具体内容，因此可以省去对日志的文本内容进行语义分析消耗的时间，提高获取训练样本的效率，进而提高故障识别模型的训练效率。

上面在介绍了设定故障对应的故障识别模型的训练方法之后，下面介绍通过上面的方式训练的到的故障识别模型进行故障识别的实现过程。需要说明的是，电子设备可以基于上述实施例训练出不同的设定故障对应的故障识别模型。

图 7 示出了本申请实施例提供的一种故障识别方法的实现流程示意图。在本申请实施例中，故障识别方法的执行主体为电子设备，例如，电脑、服务器等。电子设备可以监测 DBaaS 服务器的性能指标的相关数据，以及获取 DBaaS 服务器的实时日志。基于故障识别模型进行故障识别的电子设备和训练故障识别模型的电子设备，可以相同，也可以不同。

参照图 7，故障识别方法包括：

S701：在检测到第一性能指标的数据触发告警的情况下，确定出第一时段对应的时序数据。

其中，所述第一时段的结束时刻为所述第一性能指标对应的告警时刻；所述第一时段对应的时序数据包括实时日志对应的第一时序数据和至少一个性能指标中每个性能指标对应的第二时序数据。

电子设备基于 DBaaS 服务器在每个时刻输出的实时日志，确定出第一时序数据；基于至少一个性能指标中每个性能指标在每个时刻对应的实时监测数据，确定出每个性能指标对应的第二时序数据。

电子设备在检测到第一性能指标的数据触发告警的情况下，基于第一性能指标当前的告警时刻，从确定出的时序数据中，确定出第一时段对应的时序数据。其中，第一性能指标为至少一个性能指标中的任一性能指标。第一时段对应的时序数据包括第一时序数据和至少一个性能指标中每个性能指标对应的第二时序数据。

在实际应用中，第一时段可以为第一性能指标的告警时刻的前 15 分钟。例如，电子设备在当天上午 8 点 30 分检测到第一性能指标的数据触发告警，第一时段为 8 点 15 分到 8 点 30 分这个时段。

S702：基于至少一种设定故障中每种设定故障对应的相关性能指标，确定出至少一种第一设定故障；所述第一设定故障的相关性能指标中包含所述第一性能指标。

这里，每种设定故障对应的相关性能指标的实现方式，请参照上述确定出与设定故障相关的第一指标的相关描述，此处不赘述。

S703：将所述第一时段对应的时序数据输入所述至少一种第一设定故障中每种第一设定故障对应的故障识别模型，得到每个故障识别模型输出的识别结果。

这里，故障识别模型上述任一实施例对应的训练故障识别模型的方法训练得到。每个故障识别模型输出的识别结果为 0-1 之间的数值。识别结果表征发生的故障是故障识别模型对应的第一设定故障的概率。

S704：将所有识别结果中置信度最高的第一设定故障确定为当前发生的故障。

置信度 (confidence): 表示对应的第一设定故障可信的概率, 例如某个第一设定故障的置信度为 80%, 则表示发生的故障为第一设定故障的概率为 80%。

本实施例提供的方案中, 在检测到第一性能指标的数据触发告警的情况下, 确定出第一时段对应的时序数据, 将第一时段对应的时序数据输入至少一种第一设定故障中每种第一设定故障对应的故障识别模型, 得到每个故障识别模型输出的识别结果, 将所有识别结果中置信度最高的第一设定故障确定为当前发生的故障。基于第一时段对应的时序数据进行故障识别时, 由于第一时段对应的时序数据包括实时日志对应的第一时序数据, 第一时序数据表征基于告警时刻确定的第一时段中的每个时刻输出的日志信息量, 在确定日志信息量时不需要分析日志的具体内容, 因此可以节省对日志的文本内容进行语义分析消耗的时间, 提高故障分析效率。

图 8 示出了本申请另一实施例提供的一种故障识别方法的实现流程示意图。参照图 8, 在图 7 对应的实施例的基础上, 图 8 对应的实施例中, 故障识别方法还包括:

S705: 基于至少一台服务器中每台服务器发生的历史故障和每种历史故障对应的发生次数, 确定出至少一台服务器中所有服务器的健康评分, 以在健康评分大于或等于设定阈值的服务器中部署新的数据库实例。

电子设备基于服务器发生的历史故障的标识, 以及基于每种历史故障对应的次数, 从设定的评分表中查询对应的服务器的健康评分。其中,

电子设备中预先存储了设定的评分表, 设定的评分表包括至少一种故障等级中每个故障等级对应的所有设定故障在不同的次数范围内对应的分值。也就是说, 设定的评分表中为每个故障等级配置了对应的分值范围, 以及为每个故障等级对应的设定故障配置了不同的次数范围对应的分值, 该分值处于对应的故障等级对应的分值范围。

需要说明的是, 设定的评分表中可以按照故障等级的优先级对所有设定故障进行排序, 故障等级的优先级越高, 表征该故障等级对应的设定故障对服务器的影响程度越大。故障等级的优先级越高, 对应的分值范围越大。同一故障等级对应的设定故障发生的次数越多, 对应的分值越低。

例如, 当第一故障等级的优先级大于第二故障等级的优先级时, 第一故障等级对应的分值范围小于第二故障等级的分值范围。

针对同一故障等级对应的设定故障, 在第一次数范围大于第二次数范围时, 第一次数范围对应的分值小于第二次数范围的分值。

需要说明的是, 电子设备可以在服务器处于空闲状态的情况下, 基于服务器发生的历史故障和每种历史故障对应的次数, 确定出对应的服务器的健康评分。这样可以避免因对服务器进行健康评分而干扰服务器执行的进程, 降低服务器的数据处理效率。

本实施例提供的方案中, 基于服务器发生的历史故障和每种历史故障对应的发生次数, 确定对应的服务器的健康评分, 以在健康评分大于或等于设定阈值的服务器中部署新的数据库实例。在服务器的健康评分大于或等于设定阈值时, 表征对应的服务器存在的安全隐患较小, 在健康评分大于或等于设定阈值的服务器中部署新的数据库实例, 可以降低在部署新的数据库实例的过程中对应的服务器发生故障的概率, 还可以减少后续运维的工作量。

为实现本申请实施例的训练故障识别模型的方法, 本申请实施例还提供了一种训练故障识别模型的装置, 设置在电子设备上, 如图 9 所示, 该训练故障识别模型的装置包括:

构建单元 91, 配置为基于至少一个第一时段中每个第一时段对应的时序数据, 构建设定故障对应的至少一个正样本和至少一个负样本;

训练单元 92, 配置为基于所述设定故障对应的至少一个正样本和至少一个负样本,

对设定的二分类模型进行训练,得到所述设定故障对应的故障识别模型;其中,

所述时序数据包括第一时序数据和第二时序数据;

第一时序数据表征对应的第一时段的每个时刻输出的日志信息量;

第二时序数据表征至少一个性能指标中每个性能指标在对应的第一时段的每个时刻的监测数据。

在一实施例中,训练故障识别模型的装置还包括:

计算单元,配置为基于第一时段中每个时刻打印出的每条日志对应的第一信息量和第二信息量,计算出每条日志对应的总信息量;

确定单元,配置为基于第一时段中每个时刻打印出的每条日志对应的总信息量,得到对应的第一时段对应的第一时序数据;其中,

所述第一信息量表征日志对应的日志级别出现的第一概率;

所述第二信息量表征日志中包括的所有设定词组中每个设定词组在对应的日志级别中出现的第二概率。

在一实施例中,所述计算单元还配置为:

基于至少一个第一日志样本中每个日志级别出现的次数,确定出每个日志级别对应的第一概率;

基于所述至少一个第一日志样本中每个设定词组在每个日志级别下出现的次数,确定出每个设定词组在对应的日志级别中对应的第二概率;其中,

所述至少一个第一日志样本通过对历史日志进行采样得到。

在一实施例中,至少一个第一日志样本中每个第一日志样本均满足以下条件:

第一日志样本对应的日志级别对应的第一概率满足设定条件;

第一日志样本对应的告警类型为发生所述设定故障时监测到的告警类型;

第三概率与第四概率之间的差值小于或等于设定阈值;其中,所述第三概率表征第一日志样本对应的告警类型在所述至少一个第一日志样本中出现的概率;所述第四概率表征第一日志样本对应的告警类型在发生所述设定故障时被监测到的概率。

在一实施例中,构建单元 91 配置为:

基于第一时段中的第二时段对应的时序数据,确定出与所述设定故障相关的第一指标;其中,所述第二时段表征对应的第一时段中发生所述设定故障的时段;

基于所述第一指标在所述第二时段对应的时序数据,构建所述设定故障对应的至少一个正样本;

基于所述第一指标在所述第一时段中的第三时段对应的时序数据,构建所述设定故障对应的至少一个负样本;其中,

所述第三时段表征对应的第一时段中未发生所述设定故障的时段;

所述第一指标包括以下至少之一:

在发生所述设定故障时告警的所有性能指标中的至少一个性能指标;日志信息量。

在一实施例中,构建单元 91 配置为:

基于第二时段对应的时序数据,确定出至少两个集合;其中,所述至少两个集合包括一个第一集合和至少一个性能指标中每个性能指标对应的第二集合;所述第一集合中的元素表征相邻两个时刻输出的日志信息量之间的差值;所述第二集合中的元素表征对应的性能指标在相邻两个时刻的监测数据之间的差值;

计算出所述至少两个集合中每个集合对应的均值和标准差;

基于三西格玛准则以及基于计算出的均值和标准差,确定出所述至少两个集合中每个集合对应的第一区间;

基于所述至少两个集合中每个集合对应的第一区间和对应集合中的最大值，确定出与所述设定故障相关的第一指标；其中，所述第一指标对应的集合中的最大值未处于对应的第一区间。

在一实施例中，所述第二时段对应的时序数据未包括在第二时段内首个触发告警的第一性能指标对应的第二时序数据；构建单元 91 用于：

基于所述至少两个集合中每个集合对应的第一区间和对应集合中的最大值，确定出第二指标；其中，所述第二指标对应的集合中的最大值未处于对应的第一区间；将确定出的第二指标和所述第一性能指标确定为与所述设定故障相关的第一指标。

在一实施例中，构建单元 91 配置为：

基于所述第一指标的第一数量，确定出正样本的第二数量；

将所有第一指标在第二时段对应的时序数据以第一指标为最小单位进行全排列，得到全排列结果；

从所述全排列结果中选出第二数量的正样本；其中，

所述第一数量为大于或等于 2 的整数，所述第二数量为正整数；所述第二数量小于或等于所述第一数量对应的全排列运算结果的二分之一。

实际应用时，训练故障识别模型的装置包括的各单元可由训练故障识别模型的装置中的处理器来实现。当然，处理器需要运行存储器中存储的程序来实现上述各程序模块的功能。

需要说明的是：上述实施例提供的训练故障识别模型的装置在训练故障识别模型时，仅以上述各程序模块的划分进行举例说明，实际应用中，可以根据需要而将上述处理分配由不同的程序模块完成，即将训练故障识别模型的装置的内部结构划分成不同的程序模块，以完成以上描述的全部或者部分处理。另外，上述实施例提供的训练故障识别模型的装置与训练故障识别模型的方法实施例属于同一构思，其具体实现过程详见方法实施例，这里不再赘述。

为实现本申请实施例的故障识别方法，本申请实施例还提供了一种故障识别装置，设置在电子设备上，如图 10 所示，该故障识别装置包括：

第一确定单元 101，配置为在检测到第一性能指标的数据触发告警的情况下，确定出第一时段对应的时序数据；

第二确定单元 102，配置为基于至少一种设定故障中每种设定故障对应的相关性能指标，确定出至少一种第一设定故障；所述第一设定故障的相关性能指标中包含所述第一性能指标；

第一识别单元 103，配置为将所述第一时段对应的时序数据输入所述至少一种第一设定故障中每种第一设定故障对应的故障识别模型，得到每个故障识别模型输出的识别结果；

第二识别单元 104，配置为将所有识别结果中置信度最高的第一设定故障确定为当前发生的故障；其中，

所述第一时段的结束时刻为所述第一性能指标对应的告警时刻；

所述第一时段对应的时序数据包括实时日志对应的第一时序数据和至少一个性能指标中每个性能指标对应的第二时序数据；

所述故障识别模型基于上述任一种训练故障识别模型的方法训练得到。

在一实施例中，该故障识别装置还包括：

评分单元，配置为基于至少一台服务器中每台服务器发生的历史故障和每种历史故障对应的发生次数，确定出至少一台服务器中所有服务器的健康评分，以在健康评分大于或等于设定阈值的服务器中部署新的数据库实例。

实际应用时，故障识别装置包括的各单元可由故障识别装置中的处理器来实现。当然，处理器需要运行存储器中存储的程序实现上述各程序模块的功能。

需要说明的是：上述实施例提供的故障识别模型的装置在训练故障识别模型时，仅以上述各程序模块的划分进行举例说明，实际应用中，可以根据需要而将上述处理分配由不同的程序模块完成，即将故障识别装置的内部结构划分成不同的程序模块，以完成以上描述的全部或者部分处理。另外，上述实施例提供的故障识别装置与故障识别方法实施例属于同一构思，其具体实现过程详见方法实施例，这里不再赘述。

基于上述程序模块的硬件实现，且为了实现本申请实施例的方法，本申请实施例还提供了一种电子设备。图 11 为本申请实施例提供的电子设备的硬件组成结构示意图，

如图 11 所示，电子设备包括：

通信接口 1，能够与其它设备比如服务器等进行信息交互；

处理器 2，与通信接口 1 连接，以实现与其它设备进行信息交互，配置为运行计算机程序时，执行上述一个或多个技术方案提供的训练故障识别模型的方法；或者执行上述一个或多个技术方案提供的故障识别方法。而所述计算机程序存储在存储器 3 上。

当然，实际应用时，电子设备中的各个组件通过总线系统 4 耦合在一起。可理解，总线系统 4 用于实现这些组件之间的连接通信。总线系统 4 除包括数据总线之外，还包括电源总线、控制总线和状态信号总线。但是为了清楚说明起见，在图 11 中将各种总线都标为总线系统 4。

本申请实施例中的存储器 3 配置为存储各种类型的数据以支持电子设备的操作。这些数据的示例包括：配置为在电子设备上操作的任何计算机程序。

可以理解，存储器 3 可以是易失性存储器或非易失性存储器，也可包括易失性和非易失性存储器两者。其中，非易失性存储器可以是只读存储器（ROM，Read Only Memory）、可编程只读存储器（PROM，Programmable Read-Only Memory）、可擦除可编程只读存储器（EPROM，Erasable Programmable Read-Only Memory）、电可擦除可编程只读存储器（EEPROM，Electrically Erasable Programmable Read-Only Memory）、磁性随机存取存储器（FRAM，ferromagnetic random access memory）、快闪存储器（Flash Memory）、磁表面存储器、光盘、或只读光盘（CD-ROM，Compact Disc Read-Only Memory）；磁表面存储器可以是磁盘存储器或磁带存储器。易失性存储器可以是随机存取存储器（RAM，Random Access Memory），其用作外部高速缓存。通过示例性但不是限制性说明，许多形式的 RAM 可用，例如静态随机存取存储器（SRAM，Static Random Access Memory）、同步静态随机存取存储器（SSRAM，Synchronous Static Random Access Memory）、动态随机存取存储器（DRAM，Dynamic Random Access Memory）、同步动态随机存取存储器（SDRAM，Synchronous Dynamic Random Access Memory）、双倍数据速率同步动态随机存取存储器（DDRSDRAM，Double Data Rate Synchronous Dynamic Random Access Memory）、增强型同步动态随机存取存储器（ESDRAM，Enhanced Synchronous Dynamic Random Access Memory）、同步连接动态随机存取存储器（SLDRAM，Sync Link Dynamic Random Access Memory）、直接内存总线随机存取存储器（DDRAM，Direct Rambus Random Access Memory）。本申请实施例描述的存储器 3 旨在包括但不限于这些和任意其它适合类型的存储器。

上述本申请实施例揭示的方法可以应用于处理器 2 中，或者由处理器 2 实现。处理器 2 可能是一种集成电路芯片，具有信号的处理能力。在实现过程中，上述方法的各步骤可以通过处理器 2 中的硬件的集成逻辑电路或者软件形式的指令完成。上述的处理器 2 可以是通用处理器、DSP，或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。处理器 2 可以实现或者执行本申请实施例中的公开的各方法、步骤及

逻辑框图。通用处理器可以是微处理器或者任何常规的处理器等。结合本申请实施例所公开的方法的步骤，可以直接体现为硬件译码处理器执行完成，或者用译码处理器中的硬件及软件模块组合执行完成。软件模块可以位于存储介质中，该存储介质位于存储器 3，处理器 2 读取存储器 3 中的程序，结合其硬件完成前述方法的步骤。

5 处理器 2 执行所述程序时实现本申请实施例的各个方法中多核处理器对应的流程，为了简洁，在此不再赘述。

在示例性实施例中，本申请实施例还提供了一种存储介质，即计算机存储介质，具体为计算机可读存储介质，例如包括存储计算机程序的存储器 3，上述计算机程序可由处理器 2 执行，以完成前述图 1 至图 6 对应的实施例中的所述步骤；或者完成前述图 7
10 至图 8 对应的实施例中的所述步骤。计算机可读存储介质可以是 FRAM、ROM、PROM、EPROM、EEPROM、Flash Memory、磁表面存储器、光盘、或 CD-ROM 等存储器。

在本申请所提供的几个实施例中，应该理解到，所揭露的设备和方法，可以通过其它的方式实现。以上所描述的设备实施例仅仅是示意性的，例如，所述单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，如：多个单元或组件可以结合，或可以集成到另一个系统，或一些特征可以忽略，或不执行。
15 另外，所显示或讨论的各组成部分相互之间的耦合、或直接耦合、或通信连接可以是通过一些接口，设备或单元的间接耦合或通信连接，可以是电性的、机械的或其它形式的。

上述作为分离部件说明的单元可以是、或也可以不是物理上分开的，作为单元
20 显示的部件可以是、或也可以不是物理单元，即可以位于一个地方，也可以分布到多个网络单元上；可以根据实际的需要选择其中的部分或全部单元来实现本实施例方案的目的。

另外，在本申请各实施例中的各功能单元可以全部集成在一个处理模块中，也可以是各单元分别单独作为一个单元，也可以两个或两个以上单元集成在一个单元
25 中；上述集成的单元既可以采用硬件的形式实现，也可以采用硬件加软件功能单元的形式实现。本领域普通技术人员可以理解：实现上述方法实施例的全部或部分步骤可以通过程序指令相关的硬件来完成，前述的程序可以存储于一计算机可读存储介质中，该程序在执行时，执行包括上述方法实施例的步骤；而前述的存储介质包括：移动存储设备、只读存储器（ROM，Read-Only Memory）、随机存取存储器
30 （RAM，Random Access Memory）、磁碟或者光盘等各种可以存储程序代码的介质。

需要说明的是：“第一”、“第二”等是用于区别类似的对象，而不必用于描述特定的顺序或先后次序。

另外，本申请实施例所记载的技术方案之间，在不冲突的情况下，可以任意组合。

以上所述，仅为本申请的具体实施方式，但本申请的保护范围并不局限于此，
35 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，可轻易想到变化或替换，都应涵盖在本申请的保护范围之内。因此，本申请的保护范围应以所述权利要求要求的保护范围为准。

权利要求书

1、一种训练故障识别模型的方法，包括：

基于至少一个第一时段中每个第一时段对应的时序数据，构建设定故障对应的至少一个正样本和至少一个负样本；

5 基于所述设定故障对应的至少一个正样本和至少一个负样本，对设定的二分类模型进行训练，得到所述设定故障对应的故障识别模型；其中，

所述时序数据包括第一时序数据和第二时序数据；

第一时序数据表征对应的第一时段的每个时刻输出的日志信息量；

10 第二时序数据表征至少一个性能指标中每个性能指标在对应的第一时段的每个时刻的监测数据。

2、根据权利要求1所述的方法，其中，所述方法还包括：

基于第一时段中每个时刻打印出的每条日志对应的第一信息量和第二信息量，计算出每条日志对应的总信息量；

15 基于第一时段中每个时刻打印出的每条日志对应的总信息量，得到对应的第一时段对应的第一时序数据；其中，

所述第一信息量表征日志对应的日志级别出现的第一概率；

所述第二信息量表征日志中包括的所有设定词组中每个设定词组在对应的日志级别中出现的第二概率。

3、根据权利要求2所述的方法，其中，所述方法还包括：

20 基于至少一个第一日志样本中每个日志级别出现的次数，确定出每个日志级别对应的第一概率；

基于所述至少一个第一日志样本中每个设定词组在每个日志级别下出现的次数，确定出每个设定词组在对应的日志级别中对应的第二概率；其中，

所述至少一个第一日志样本通过对历史日志进行采样得到。

25 4、根据权利要求3所述的方法，其中，至少一个第一日志样本中每个第一日志样本均满足以下条件：

第一日志样本对应的日志级别对应的第一概率满足设定条件；

第一日志样本对应的告警类型为发生所述设定故障时监测到的告警类型；

30 第三概率与第四概率之间的差值小于或等于设定阈值；其中，所述第三概率表征第一日志样本对应的告警类型在所述至少一个第一日志样本中出现的概率；所述第四概率表征第一日志样本对应的告警类型在发生所述设定故障时被监测到的概率。

5、根据权利要求1-4任一项所述的方法，其中，所述基于至少一个第一时段对应的时序数据，构建设定故障对应的至少一个正样本和至少一个负样本，包括：

35 基于第一时段中的第二时段对应的时序数据，确定出与所述设定故障相关的第一指标；其中，所述第二时段表征对应的第一时段中发生所述设定故障的时段；

基于所述第一指标在所述第二时段对应的时序数据，构建所述设定故障对应的至少一个正样本；

基于所述第一指标在所述第一时段中的第三时段对应的时序数据，构建所述设定故障对应的至少一个负样本；其中，

40 所述第三时段表征对应的第一时段中未发生所述设定故障的时段；

所述第一指标包括以下至少之一：

在发生所述设定故障时告警的所有性能指标中的至少一个性能指标；

日志信息量。

6、根据权利要求 5 所述的方法，其中，所述基于第一时段中的第二时段对应的时序数据，确定出与所述设定故障相关的第一指标，包括：

基于第二时段对应的时序数据，确定出至少两个集合；其中，所述至少两个集合包括一个第一集合和至少一个性能指标中每个性能指标对应的第二集合；所述第一集合中的元素表征相邻两个时刻输出的日志信息量之间的差值；所述第二集合中的元素表征对应的性能指标在相邻两个时刻的监测数据之间的差值；

计算出所述至少两个集合中每个集合对应的均值和标准差；

基于三西格玛准则以及基于计算出的均值和标准差，确定出所述至少两个集合中每个集合对应的第一区间；

基于所述至少两个集合中每个集合对应的第一区间和对应集合中的最大值，确定出与所述设定故障相关的第一指标；其中，所述第一指标对应的集合中的最大值未处于对应的第一区间。

7、根据权利要求 6 所述的方法，其中，所述第二时段对应的时序数据未包括在第二时段内首个触发告警的第一性能指标对应的第二时序数据；

所述基于所述至少两个集合中每个集合对应的第一区间和对应集合中的最大值，确定出与所述设定故障相关的第一指标，包括：

基于所述至少两个集合中每个集合对应的第一区间和对应集合中的最大值，确定出第二指标；其中，所述第二指标对应的集合中的最大值未处于对应的第一区间；

将确定出的第二指标和所述第一性能指标确定为与所述设定故障相关的第一指标。

8、根据权利要求 5 所述的方法，其中，所述基于所述第一指标在所述第二时段对应的时序数据，构建所述设定故障对应的至少一个正样本，包括：

基于所述第一指标的第一数量，确定出正样本的第二数量；

将所有第一指标在第二时段对应的时序数据以第一指标为最小单位进行全排列，得到全排列结果；

从所述全排列结果中选出第二数量的正样本；其中，

所述第一数量为大于或等于 2 的整数，所述第二数量为正整数；

所述第二数量小于或等于所述第一数量对应的全排列运算结果的二分之一。

9、一种故障识别方法，包括：

在检测到第一性能指标的数据触发告警的情况下，确定出第一时段对应的时序数据；

基于至少一种设定故障中每种设定故障对应的相关性能指标，确定出至少一种第一设定故障；所述第一设定故障的相关性能指标中包含所述第一性能指标；

将所述第一时段对应的时序数据输入所述至少一种第一设定故障中每种第一设定故障对应的故障识别模型，得到每个故障识别模型输出的识别结果；

将所有识别结果中置信度最高的第一设定故障确定为当前发生的故障；其中，

所述第一时段的结束时刻为所述第一性能指标对应的告警时刻；

所述第一时段对应的时序数据包括实时日志对应的第一时序数据和至少一个性能指标中每个性能指标对应的第二时序数据；

所述故障识别模型基于权利要求 1 至 8 任一项所述的方法训练得到。

10、根据权利要求 9 所述的方法，其中，所述方法还包括：

基于至少一台服务器中每台服务器发生的历史故障和每种历史故障对应的发生次数，确定出至少一台服务器中所有服务器的健康评分，以在健康评分大于或等于设定阈值的服务器中部署新的数据库实例。

11、一种训练故障识别模型的装置，包括：

构建单元，配置为基于至少一个第一时段中每个第一时段对应的时序数据，构建设定故障对应的至少一个正样本和至少一个负样本；

5 训练单元，配置为基于所述设定故障对应的至少一个正样本和至少一个负样本，对设定的二分类模型进行训练，得到所述设定故障对应的故障识别模型；其中，

所述时序数据包括第一时序数据和第二时序数据；

第一时序数据表征对应的第一时段的每个时刻输出的日志信息量；

第二时序数据表征至少一个性能指标中每个性能指标在对应的第一时段的每个时刻的监测数据。

10 12、一种故障识别装置，包括：

第一确定单元，配置为在检测到第一性能指标的数据触发告警的情况下，确定出第一时段对应的时序数据；

15 第二确定单元，配置为基于至少一种设定故障中每种设定故障对应的相关性能指标，确定出至少一种第一设定故障；所述第一设定故障的相关性能指标中包含所述第一性能指标；

第一识别单元，配置为将所述第一时段对应的时序数据输入所述至少一种第一设定故障中每种第一设定故障对应的故障识别模型，得到每个故障识别模型输出的识别结果；

20 第二识别单元，配置为将所有识别结果中置信度最高的第一设定故障确定为当前发生的故障；其中，

所述第一时段的结束时刻为所述第一性能指标对应的告警时刻；

所述第一时段对应的时序数据包括实时日志对应的第一时序数据和至少一个性能指标中每个性能指标对应的第二时序数据；

所述故障识别模型基于权利要求 1 至 8 任一项所述的方法训练得到。

25 13、一种电子设备，包括：处理器和配置为存储能够在处理器上运行的计算机程序的存储器，

其中，所述处理器配置为运行所述计算机程序时，执行以下至少一项：

权利要求 1 至 8 任一项所述的方法的步骤；

权利要求 9 至 10 任一项所述的方法的步骤。

30 14、一种存储介质，其上存储有计算机程序，所述计算机程序被处理器执行时实现以下至少一项：

权利要求 1 至 8 任一项所述的方法的步骤；

权利要求 9 至 10 任一项所述的方法的步骤。

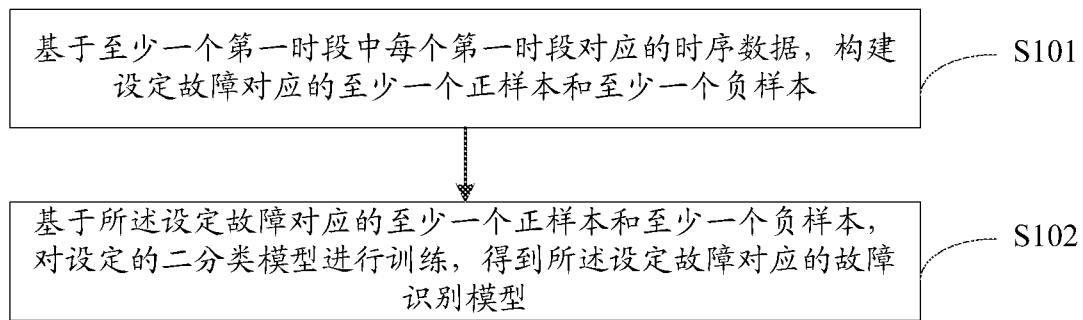


图 1

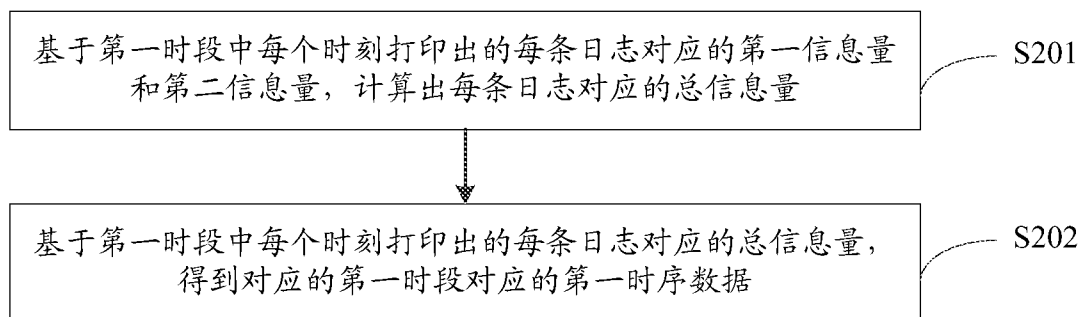


图 2

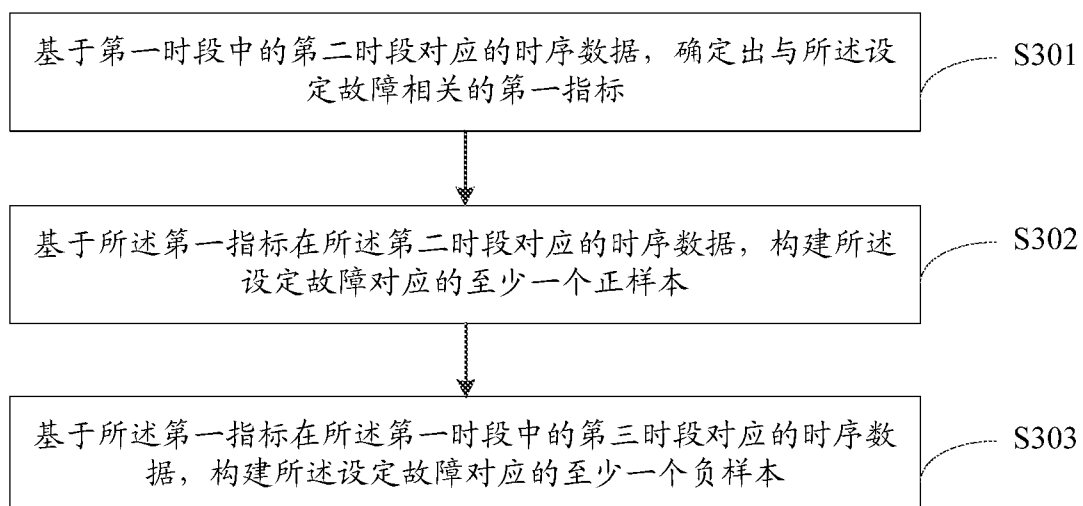


图 3

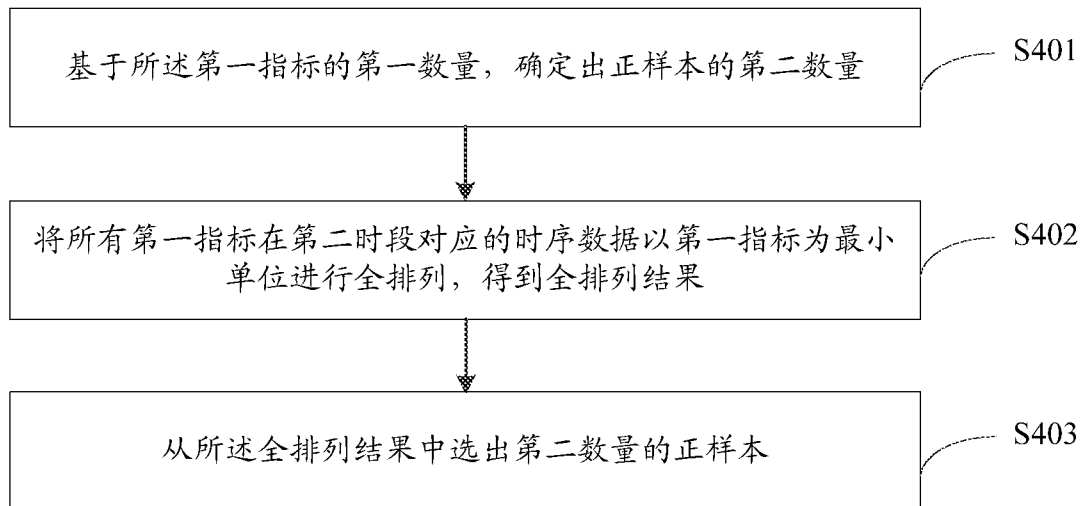


图 4

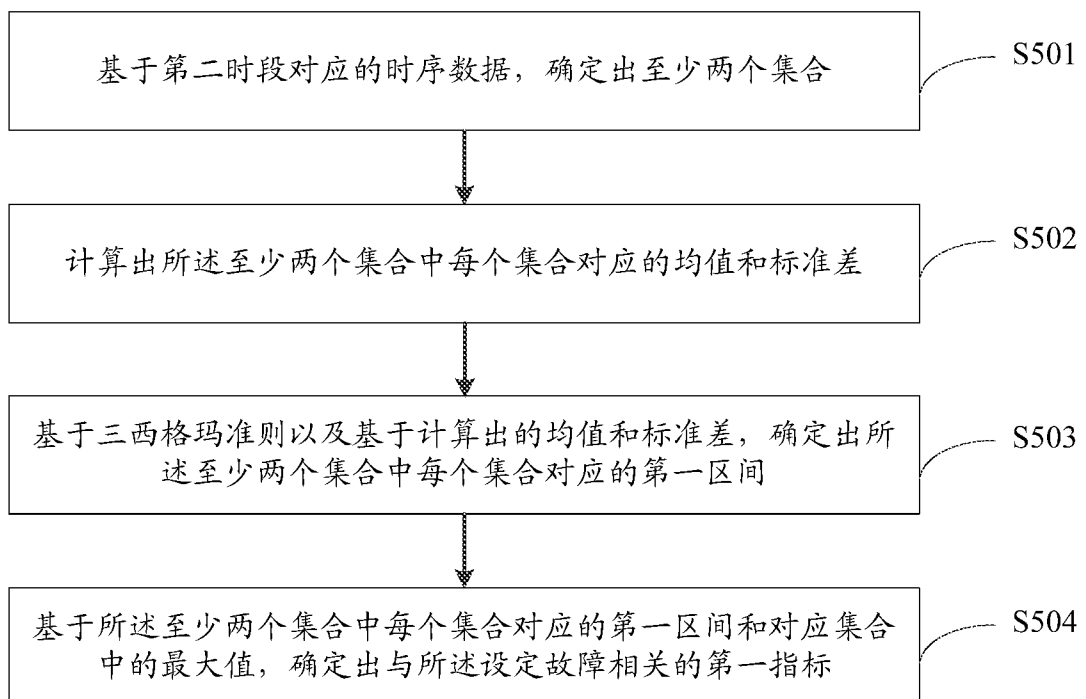


图 5

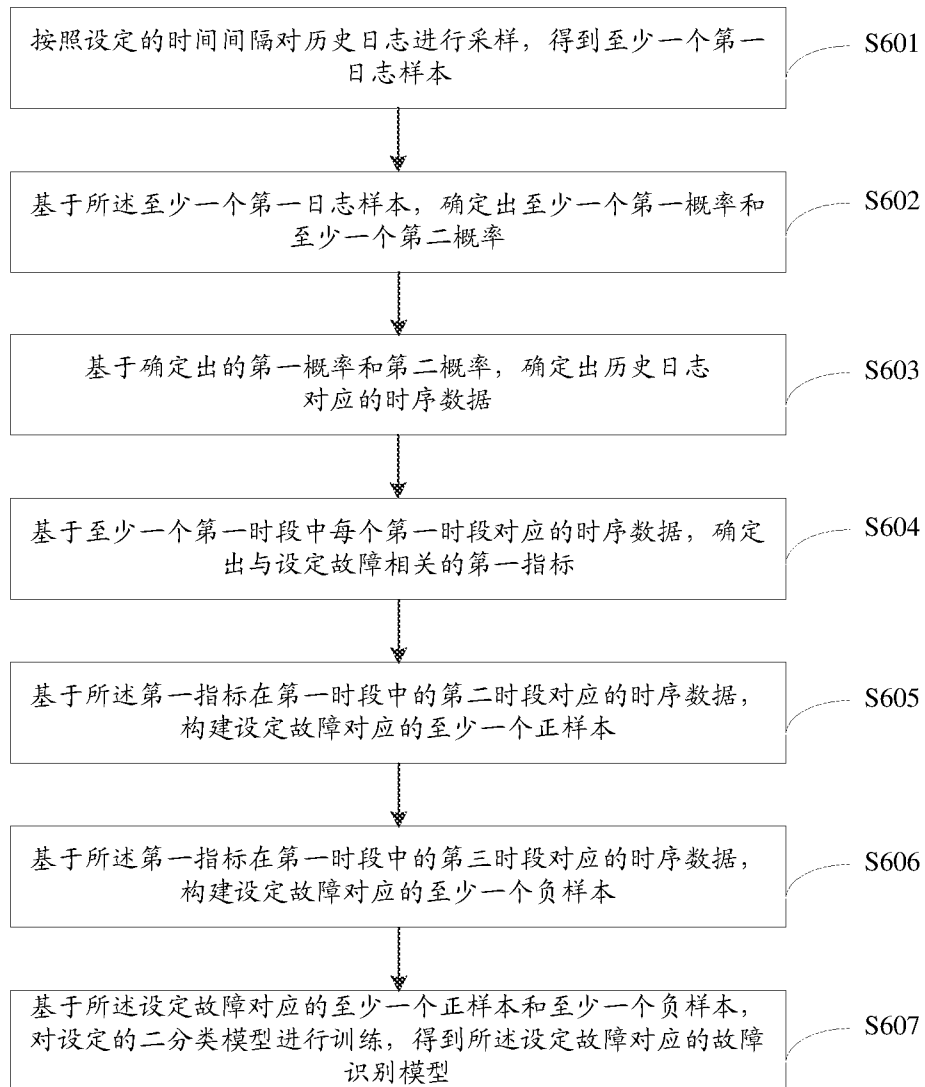


图 6

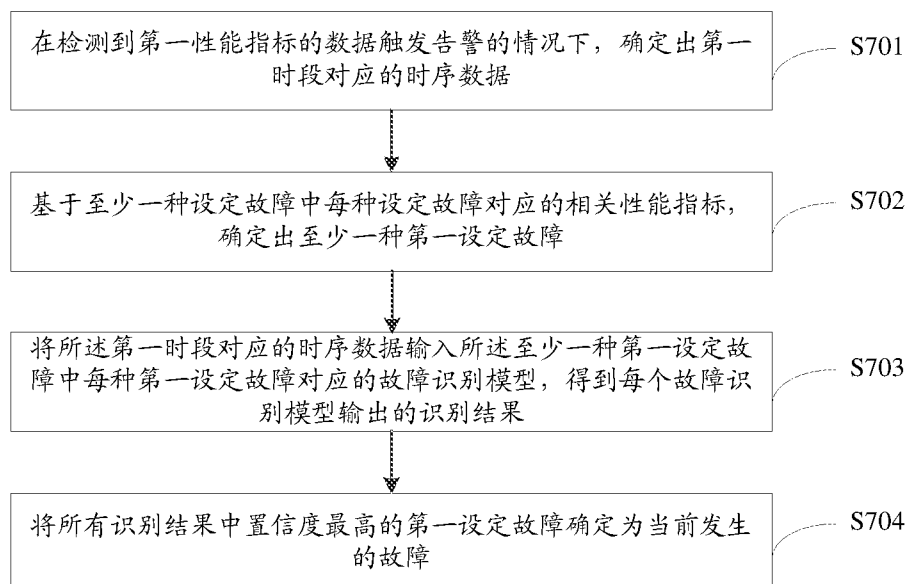


图 7

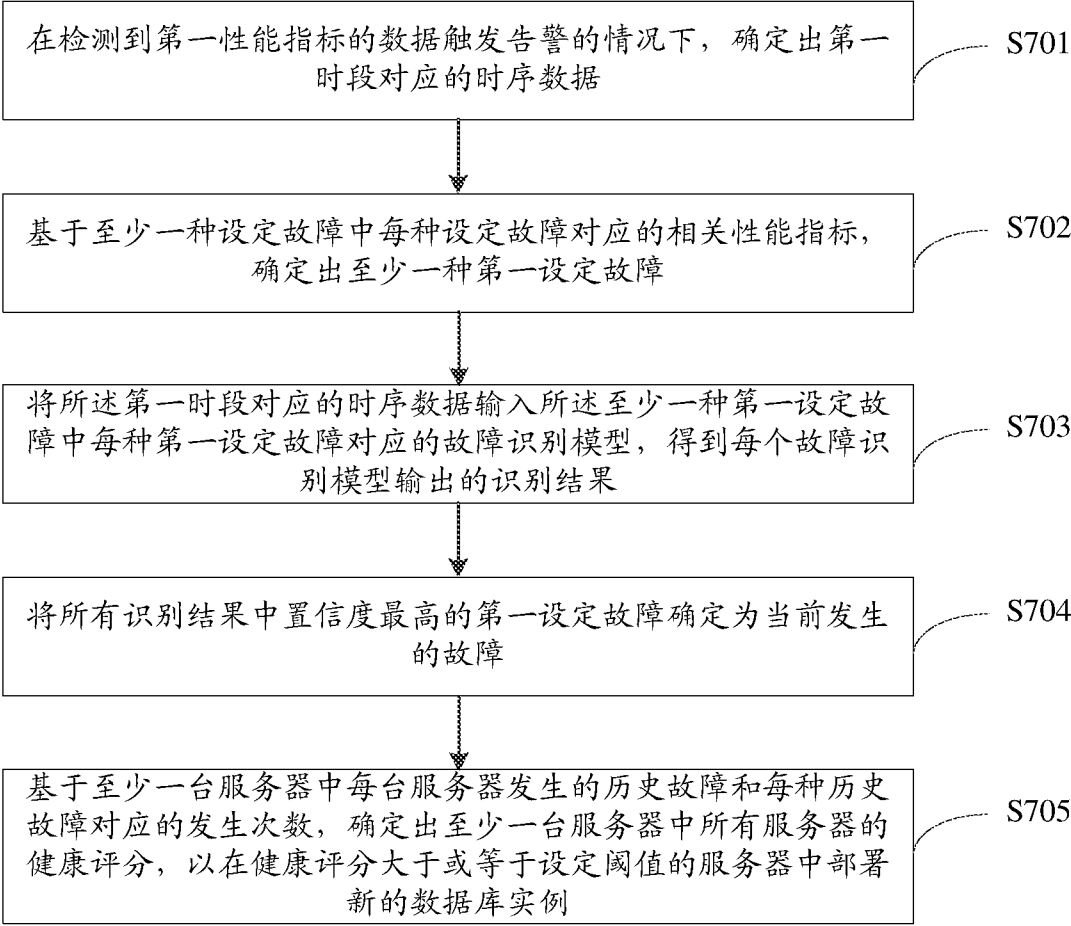


图 8

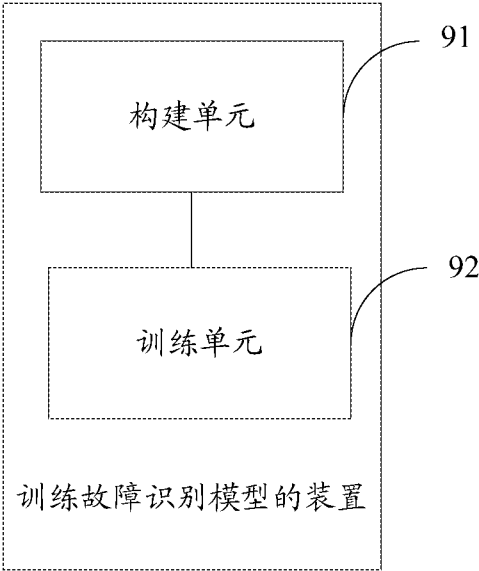


图 9

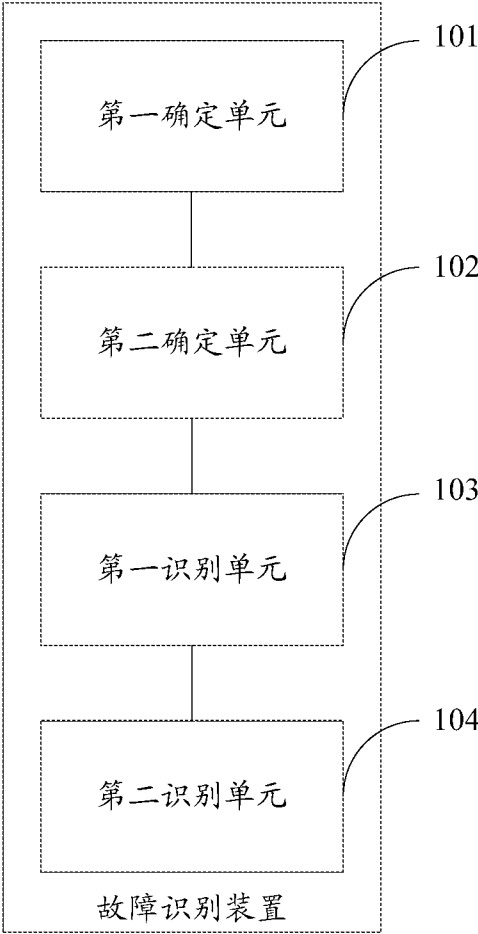


图 10

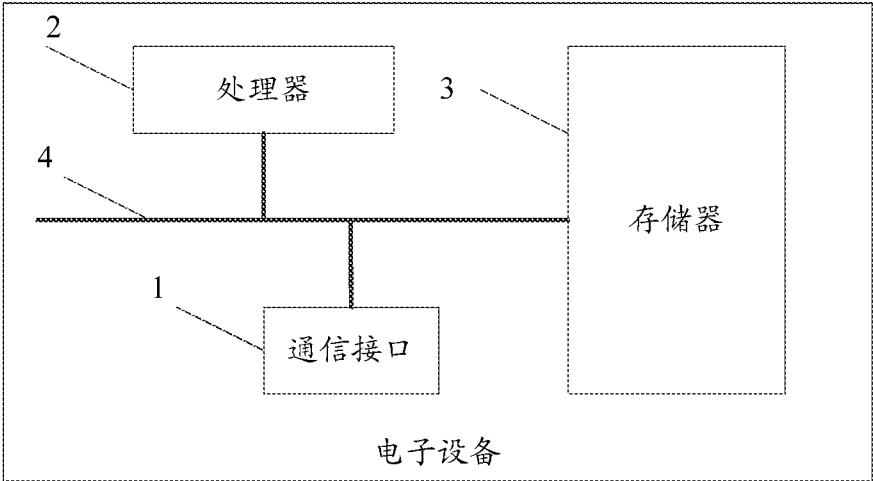


图 11

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/123363

A. CLASSIFICATION OF SUBJECT MATTER G06K 9/62(2022.01)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) G06K Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) CNKI, CNABS, DWPI, SIPOABS, CNTXT: 故障识别, 正样本, 负样本, 时序, 日志, 时段, fault identification, positive sample, negative sample, time sequence, blog, period																				
C. DOCUMENTS CONSIDERED TO BE RELEVANT <table border="1"> <thead> <tr> <th>Category*</th> <th>Citation of document, with indication, where appropriate, of the relevant passages</th> <th>Relevant to claim No.</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 112308126 A (SHENZHEN QIANHAI WEBANK CO., LTD.) 02 February 2021 (2021-02-02) description paragraphs [0001]-[0309]</td> <td>1-14</td> </tr> <tr> <td>X</td> <td>CN 111585799 A (HANGZHOU DPTECH TECHNOLOGIES CO., LTD.) 25 August 2020 (2020-08-25) description, paragraphs 39-95</td> <td>1, 11, 13, 14</td> </tr> <tr> <td>Y</td> <td>CN 111585799 A (HANGZHOU DPTECH TECHNOLOGIES CO., LTD.) 25 August 2020 (2020-08-25) description, paragraphs 39-95</td> <td>9, 12-14</td> </tr> <tr> <td>Y</td> <td>CN 109639450 A (PING AN YIQIANBAO E-COMMERCE CO., LTD.) 16 April 2019 (2019-04-16) description, paragraphs 5-22</td> <td>9, 12-14</td> </tr> <tr> <td>A</td> <td>EP 3336636 A1 (PALANTIR TECHNOLOGIES INC.) 20 June 2018 (2018-06-20) entire document</td> <td>1-14</td> </tr> </tbody> </table>			Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.	PX	CN 112308126 A (SHENZHEN QIANHAI WEBANK CO., LTD.) 02 February 2021 (2021-02-02) description paragraphs [0001]-[0309]	1-14	X	CN 111585799 A (HANGZHOU DPTECH TECHNOLOGIES CO., LTD.) 25 August 2020 (2020-08-25) description, paragraphs 39-95	1, 11, 13, 14	Y	CN 111585799 A (HANGZHOU DPTECH TECHNOLOGIES CO., LTD.) 25 August 2020 (2020-08-25) description, paragraphs 39-95	9, 12-14	Y	CN 109639450 A (PING AN YIQIANBAO E-COMMERCE CO., LTD.) 16 April 2019 (2019-04-16) description, paragraphs 5-22	9, 12-14	A	EP 3336636 A1 (PALANTIR TECHNOLOGIES INC.) 20 June 2018 (2018-06-20) entire document	1-14
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.																		
PX	CN 112308126 A (SHENZHEN QIANHAI WEBANK CO., LTD.) 02 February 2021 (2021-02-02) description paragraphs [0001]-[0309]	1-14																		
X	CN 111585799 A (HANGZHOU DPTECH TECHNOLOGIES CO., LTD.) 25 August 2020 (2020-08-25) description, paragraphs 39-95	1, 11, 13, 14																		
Y	CN 111585799 A (HANGZHOU DPTECH TECHNOLOGIES CO., LTD.) 25 August 2020 (2020-08-25) description, paragraphs 39-95	9, 12-14																		
Y	CN 109639450 A (PING AN YIQIANBAO E-COMMERCE CO., LTD.) 16 April 2019 (2019-04-16) description, paragraphs 5-22	9, 12-14																		
A	EP 3336636 A1 (PALANTIR TECHNOLOGIES INC.) 20 June 2018 (2018-06-20) entire document	1-14																		
☐ 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 11 January 2022		Date of mailing of the international search report 18 January 2022																		
Name and mailing address of the ISA/CN China National Intellectual Property Administration (ISA/CN) No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088, China Facsimile No. (86-10)62019451		Authorized officer Telephone No.																		

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2021/123363

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	112308126	A	02 February 2021	None			
CN	111585799	A	25 August 2020	None			
CN	109639450	A	16 April 2019	None			
EP	3336636	A1	20 June 2018	US	2019287014	A1	19 September 2019

国际检索报告

国际申请号

PCT/CN2021/123363

A. 主题的分类

G06K 9/62 (2022.01) 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)

G06K

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用))

CNKI, CNABS, DWPI, SIPOABS, CNTXT: 故障识别, 正样本, 负样本, 时序, 日志, 时段, fault identification, positive sample, negative sample, time sequence, blog, period

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 112308126 A (深圳前海微众银行股份有限公司) 2021年2月2日 (2021 - 02 - 02) 说明书第[0001]-[0309]段	1-14
X	CN 111585799 A (杭州迪普科技股份有限公司) 2020年8月25日 (2020 - 08 - 25) 说明书第39-95段	1、11、13-14
Y	CN 111585799 A (杭州迪普科技股份有限公司) 2020年8月25日 (2020 - 08 - 25) 说明书第39-95段	9, 12-14
Y	CN 109639450 A (平安壹钱包电子商务有限公司) 2019年4月16日 (2019 - 04 - 16) 说明书第5-22段	9, 12-14
A	EP 3336636 A1 (PALANTIR TECHNOLOGIES INC.) 2018年6月20日 (2018 - 06 - 20) 全文	1-14

☐ 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

2022年1月11日

国际检索报告邮寄日期

2022年1月18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

刘瑛

电话号码 010-62412234

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2021/123363

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	112308126	A	2021年2月2日	无			
CN	111585799	A	2020年8月25日	无			
CN	109639450	A	2019年4月16日	无			
EP	3336636	A1	2018年6月20日	US	2019287014	A1	2019年9月19日