

(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
2021 年 9 月 30 日 (30.09.2021)



(10) 国际公布号
WO 2021/189844 A1

- (51) 国际专利分类号:
G06F 16/2458 (2019.01) **G06N 3/04** (2006.01)
- (21) 国际申请号: PCT/CN2020/125003
- (22) 国际申请日: 2020 年 10 月 30 日 (30.10.2020)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202011004044.3 2020年9月22日 (22.09.2020) CN
- (71) 申请人: 平安科技(深圳)有限公司(PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN]; 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。
- (72) 发明人: 邓悦(DENG, Yue); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。 郑

立颖(ZHENG, Liying); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。 徐亮(XU, Liang); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(74) 代理人: 深圳市沃德知识产权代理事务所(普通合伙)(SHENZHEN WORLD INTELLECTUAL PROPERTY AGENCY (GENERAL PARTNERSHIP)); 中国广东省深圳市福田区梅林街道中康路 128 号卓越城 1 期 2 栋 301B 于志光, Guangdong 518000 (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,

(54) Title: DETECTION METHOD AND APPARATUS FOR MULTIVARIATE KPI TIME SERIES, AND DEVICE AND STORAGE MEDIUM

(54) 发明名称: 多元KPI时间序列的检测方法、装置、设备及存储介质

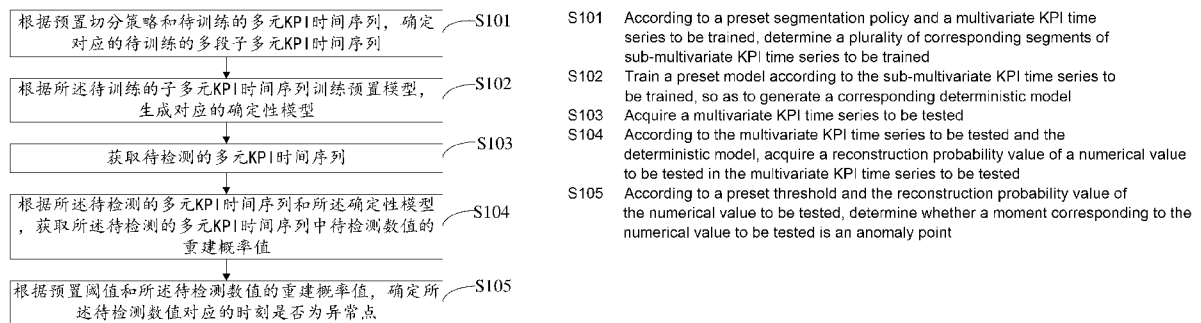


图 1

(57) Abstract: A detection method and apparatus for a multivariate KPI time series, and a computer device and a computer-readable storage medium, which relate to the technical field of artificial intelligence. The method comprises: according to a preset segmentation policy and a multivariate KPI time series to be trained, determining a plurality of corresponding segments of sub-multivariate KPI time series to be trained (S101), wherein the sub-multivariate KPI time series to be trained are untagged; training a preset model according to the sub-multivariate KPI time series to be trained, so as to generate a corresponding deterministic model (S102); acquiring a multivariate KPI time series to be tested (S103); according to the multivariate KPI time series to be tested and the deterministic model, acquiring a reconstruction probability value of a numerical value to be tested in the multivariate KPI time series to be tested (S104); and according to a preset threshold and the reconstruction probability value of the numerical value to be tested, determining whether a moment corresponding to the numerical value to be tested is an anomaly point (S105). The deterministic model is generated by means of the untagged multivariate KPI time series, the robustness of the model is enhanced, and the detection accuracy is improved.

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,
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))。
- 在修改权利要求的期限届满之前进行, 在收到该
修改后将重新公布 (细则48.2(h))。
- 根据申请人的请求, 在条约第21条(2)(a)所规定的
期限届满之前进行。

(57) 摘要: 一种多元KPI时间序列的检测方法、装置、计算机设备及计算机可读存储介质, 涉及人工智能技术领域, 该方法包括: 根据预置切分策略和待训练的多元KPI时间序列, 确定对应的待训练的多段子多元KPI时间序列 (S101), 其中, 待训练的子多元KPI时间序列为无标签; 根据所述待训练的子多元KPI时间序列训练预置模型, 生成对应的确定性模型 (S102); 获取待检测的多元KPI时间序列 (S103); 根据所述待检测的多元KPI时间序列和确定性模型, 获取所述待检测的多元KPI时间序列中待检测数值的重建概率值 (S104); 根据预置阈值和所述待检测数值的重建概率值, 确定所述待检测数值对应的时刻是否为异常点 (S105), 实现通过无标签的多元KPI时间序列生成确定性模型, 增强了模型的鲁棒性, 并提高了检测的准确率。

多元 KPI 时间序列的检测方法、装置、设备及存储介质

本申请要求于2020年9月22日提交中国专利局、申请号为CN2020110040443、名称为“多元KPI时间序列的检测方法、装置、设备及存储介质”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

本申请涉及人工智能技术领域，尤其涉及一种多元KPI时间序列的检测方法、装置、计算机设备及计算机可读存储介质。

背景技术

工业设备，例如服务器、航天器、机器人辅助系统、引擎等通常会产生多个时间序列指标，以便可以及时检测到每个设备的行为异常。来自同一设备或同一个实体的多个单变量时间序列形成一个多元KPI时间序列。通常，直接使用多元KPI时间序列在实体级别检测实体异常的效果会优于多个单变量时间序列的异常检测效果。

多元KPI时间序列的异常检测主要分为基于确定性模型和基于随机性模型，基于确定性模型的主要是重建“正常”时间序列行为，并使用重建误差进行多传感器异常检测。发明人意识到，确定性模型需要大量的带标签时间序列进行训练；而基于随机性模型获取历史多元KPI时间序列的数值范围，通过该数值范围检测当前多元KPI时间序列是否异常，检测的准确率较低。

发明内容

本申请提供一种多元 KPI 时间序列的检测方法，所述多元 KPI 时间序列的检测方法包括以下步骤：

根据预置切分策略和待训练的多元 KPI 时间序列，确定对应的待训练的多段子多元 KPI 时间序列，其中，所述待训练的子多元 KPI 时间序列为无标签；

根据所述待训练的子多元 KPI 时间序列训练预置模型，生成对应的确定性模型；

获取待检测的多元 KPI 时间序列；

根据所述待检测的多元 KPI 时间序列和所述确定性模型，获取所述待检测的多元 KPI 时间序列中待检测数值的重建概率值；

根据预置阈值和所述待检测数值的重建概率值，确定所述待检测数值对应的时刻是否为异常点。

本申请还提供一种多元 KPI 时间序列的检测装置，所述多元 KPI 时间序列的检测装置包括：

得到模块，用于根据预置切分策略和待训练的多元 KPI 时间序列，得到对应的待训练的多段子多元 KPI 时间序列，其中，所述待训练的子多元 KPI 时间序列为无标签；

生成模块，用于根据所述待训练的子多元 KPI 时间序列训练预置模型，生成对应的确定性模型；

第一获取模块，用于获取待检测的多元 KPI 时间序列；

第二获取模块，用于根据所述待检测的多元 KPI 时间序列和所述确定性模型，获取所述待检测的多元 KPI 时间序列中待检测数值的重建概率值；

确定模块，用于根据预置阈值和所述待检测数值的重建概率值，确定所述待检测数值对应的时刻是否为异常点。

本申请还提供一种计算机设备，所述计算机设备包括处理器、存储器、以及存储在所述存储器上并可被所述处理器执行的计算机程序，其中所述计算机程序被所述处理器执行

时，实现如下步骤：

根据预置切分策略和待训练的多元 KPI 时间序列，确定对应的待训练的多段子多元 KPI 时间序列，其中，所述待训练的子多元 KPI 时间序列为无标签；

根据所述待训练的子多元 KPI 时间序列训练预置模型，生成对应的确定性模型；

获取待检测的多元 KPI 时间序列；

根据所述待检测的多元 KPI 时间序列和所述确定性模型，获取所述待检测的多元 KPI 时间序列中待检测数值的重建概率值；

根据预置阈值和所述待检测数值的重建概率值，确定所述待检测数值对应的时刻是否为异常点。

本申请还提供一种计算机可读存储介质，所述计算机可读存储介质上存储有计算机程序，其中所述计算机程序被处理器执行时，实现如下步骤：

根据预置切分策略和待训练的多元 KPI 时间序列，确定对应的待训练的多段子多元 KPI 时间序列，其中，所述待训练的子多元 KPI 时间序列为无标签；

根据所述待训练的子多元 KPI 时间序列训练预置模型，生成对应的确定性模型；

获取待检测的多元 KPI 时间序列；

根据所述待检测的多元 KPI 时间序列和所述确定性模型，获取所述待检测的多元 KPI 时间序列中待检测数值的重建概率值；

根据预置阈值和所述待检测数值的重建概率值，确定所述待检测数值对应的时刻是否为异常点。

附图说明

为了更清楚地说明本申请实施例技术方案，下面将对实施例描述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他的附图。

图1为本申请实施例提供的一种多元KPI时间序列的检测方法的流程示意图；

图2为图1中的多元KPI时间序列的检测方法的子步骤流程示意图；

图3为本申请实施例中确定性模型的解码和编码示意图；

图4为本申请实施例提供的一种多元KPI时间序列的检测装置的示意性框图；

图5为本申请一实施例涉及的计算机设备的结构示意图。

本申请目的的实现、功能特点及优点将结合实施例，参照附图做进一步说明。

具体实施方式

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

附图中所示的流程图仅是示例说明，不是必须包括所有的内容和操作/步骤，也不是必须按所描述的顺序执行。例如，有的操作/步骤还可以分解、组合或部分合并，因此实际执行的顺序有可能根据实际情况改变。

本申请实施例提供一种多元 KPI 时间序列的检测方法、装置、计算机设备及计算机可读存储介质。其中，该多元 KPI 时间序列的检测方法可应用于计算机设备中，该计算机设备可以是笔记本电脑、台式电脑等电子设备。

下面结合附图，对本申请的一些实施方式作详细说明。在不冲突的情况下，下述的实施例及实施例中的特征可以相互组合。

请参照图 1，图 1 为本申请的实施例提供的一种多元 KPI 时间序列的检测方法的流程

示意图。

如图 1 所示, 该多元 KPI 时间序列的检测方法包括步骤 S101 至步骤 S105。

步骤 S101、根据预置切分策略和待训练的多元 KPI 时间序列, 确定对应的待训练的多段子多元 KPI 时间序列, 其中, 所述待训练的子多元 KPI 时间序列为无标签。

示范性的, 获取待训练多元 KPI 时间序列, 通过预置切分策略对获取到的待训练多元 KPI 时间序列进行数据处理, 预置切分策略包括 $T+1$ 的滑动窗口信息, 其中, T 可以为 0。例如, 当 T 为 1 时, 则滑动窗口的长度为 2 个时刻。当获取到的待训练多元 KPI 时间序列包括 0-24 个时刻, 则基于滑动窗口的长度为 2 个时刻对待训练多元 KPI 时间序列的 0-24 个时刻进行切分, 得到 23 段子多元 KPI 时间序列。或者, 当 T 为 0 时, 则滑动窗口的长度为 1 个时刻。当获取到的待训练多元 KPI 时间序列包括 0-24 个时刻, 则基于滑动窗口的长度为 1 个时刻对待训练多元 KPI 时间序列的 0-24 个时刻进行切分, 得到 24 段子多元 KPI 时间序列。其中, 每一段子多元 KPI 时间序列上时刻对应有个数值。

步骤 S102、根据所述待训练的子多元 KPI 时间序列训练预置模型, 生成对应的确定性模型。

示范性的, 通过子多元 KPI 时间序列对预置模型进行训练, 生成对应的确定性模型, 其中, 该子多元 KPI 时间序列为不带标签的子多元 KPI 时间序列。例如, 将该子多元 KPI 时间序列输入预置模型, 通过预置模型中的网络层提取该子多元 KPI 时间序列中各个时刻对应的数值, 获取各个时刻对应的数值之间的向量特征, 基于该各个时刻对应的数值之间的向量特征对该网络层的权重参数进行训练, 生成对应的确定性模型。或者, 将该子多元 KPI 时间序列输入预置模型, 为每个输入子多元 KPI 时间序列的数值样本建立特定分布以获取辅助变量, 再根据辅助变量重建原始数值样本, 通过最大化损失函数以改进预置模型的网络参数, 生成对应的确定性模型。

在一实施例中, 具体地, 参照图 2, 步骤 S102 包括: 子步骤 S1021 至子步骤 S1024。

子步骤 S1021、将所述待训练的子多元 KPI 时间序列输入所述预置模型, 提取所述待训练的子多元 KPI 时间序列中的数值。

将该子多元 KPI 时间序列输入预置模型, 通过预置模型的输入层提取该子多元 KPI 时间序列中数值, 其中, 该子多元 KPI 时间序列包括多个时刻, 且每个时刻对应多个数值。该数值可以为第一时刻的第一数值也可以为第一时刻的第二数值, 或者, 为第二时刻的第一数值等。

子步骤 S1022、根据预置编码程序对所述数值进行编码, 获取所述数值的第一高斯分布和第一辅助变量。

预置模型包括编码程序, 通过该编码程序对该数值进行编码, 得到该数值编码后的第一高斯分布, 从该第一高斯分布中随机抽取一个对角高斯向量, 将该对角高斯向量作为该数值的第一辅助变量。

在一实施例中, 所述预置编码程序包括第一预置神经网络模型和第一全连接层, 所述根据预置编码程序对所述数值进行编码, 获取所述数值的第一高斯分布和第一辅助变量, 包括: 根据所述数值和第一预置神经网络模型, 得到所述第一预置神经网络模型的第一隐藏变量; 根据第一预置全连接层和所述第一隐藏变量, 得到所述数值对应的第一高斯分布; 基于所述第一高斯分布得到所述数值的第一辅助变量。

示范性的, 如图 3 所示, 预置模型包括编码程序和解码程序, 其中, 该编码程序包括第一预置神经网络 GPU 和第一全连接层 h_1 ; 解码程序包括第二预置神经网络 GPU 和第二全连接层 h_2 。在获取到子多元 KPI 时间序列中的数值时, 获取第一预置神经网络 GPU 之前产生的第一隐藏变量, 通过将该数值和第一隐藏变量输入该第一预置神经网络 GPU 中, 得到第一预置神经网络 GPU 输出的第二隐藏变量。例如, 获取到子多元 KPI 时间序列中的数值 x_t , 获取第一预置神经网络 GPU 之前产生的第一隐藏变量 e_{t-1} , 将该 x_t 和 e_{t-1} 输入第一预置

神经网络 GPU，通过该第一预置神经网络 GPU 中第一预置公式 $e_t = (1 - c_t^e) \times \tanh[w^e x_t + u^e (r_t^e \times e_{t-1}) + b^e] + c_t^e \times e_{t-1}$ ，其中， w^e 、 u^e 、 b^e 为变量参数，通过获取第一神经网络或预置模型的参数矩阵，该参数矩阵是以 0 为中心分布的，随机抽取该参数矩阵中的参数作为 w^e 、 u^e 、 b^e 的初始参数。 c_t^e 是 GRU 中的更新门，用于决定如何将新输入的自变量与之前的时序信息结合起来。 r_t^e 是 GRU 中的重置门，用于决定保留多少此前的时序信息。通过第二预置公式得到 $r_t^e = \text{sigmoid}(w^{r^e} x_t + u^{r^e} e_{t-1} + b^{r^e})$ ，其中， w^{r^e} 、 u^{r^e} 、 b^{r^e} 为为变量参数，通过获取第一神经网络或预置模型的参数矩阵，该参数矩阵是以 0 为中心分布的，随机抽取该参数矩阵中的参数作为 w^{r^e} 、 u^{r^e} 、 b^{r^e} 的初始参数。通过第三预置公式得到 $c_t^e = \text{sigmoid}(w^{c^e} x_t + u^{c^e} e_{t-1} + b^{c^e})$ ，其中， w^{c^e} 、 u^{c^e} 、 b^{c^e} 为为变量参数，通过获取第一神经网络或预置模型的参数矩阵，该参数矩阵是以 0 为中心分布的，随机抽取该参数矩阵中的参数作为 w^{c^e} 、 u^{c^e} 、 b^{c^e} 的初始参数。

在得到该数值的第一隐藏变量时，获取该第一全连接层之前的第二隐藏变量，将得到的第一隐藏变量与的第二隐藏变量连接，得到向量信息。将该向量信息输入第一全连接层，通过该第一全连接层中的网络层以预置 ReLU 作为激活函数，分别经过下面两个线性变换和 *softplus* 激活函数变换得到均值和标准差。例如，在得到该数值的第一隐藏变量 e_t 以及该第一全连接层之前的第二隐藏变量 z_{t-1} ，将第一隐藏变量 e_t 与第二隐藏变量 z_{t-1} 连接，得到隐藏变量向量 (z_{t-1}, e_t) ，通过第一全连接层的第一线性变换函数 $u_{z_t} = w^{u_z} h_1(z_{t-1}, e_t) + b^{u_z}$ ，得到该数值的第一均值。其中， $w^{u_z} h_1$ 、 b^{u_z} 为常数。通过第一全连接层的第二线性变换函数 $\sigma_{z_t} = \text{softplus}[w^{u_z} h_1(z_{t-1}, e_t) + b^{u_z}] + \varepsilon^{\sigma_z}$ ，得到该数值的第一标准差。其中， $w^{u_z} h_1$ 、 b^{u_z} 、 ε^{σ_z} 为常数， ε 是一个很小的常数，为了防止计算时数值溢出而设置。

在得到该数值的第一均值和第一标准差时，通过第一均值和第一标准差构建该数值的第一高斯分布。例如，在获取第一均值 u_{z_t} 和第一标准差 σ_{z_t} 时，构建该数值对应时刻的特定高斯分布 $N(u_{z_t}, \sigma_{z_t}^2 I)$ 。在构建该数值的第一高斯分布时，随机从该第一高斯分布抽取一个对角高斯向量作为该数值的第一辅助变量。例如，从该特定高斯分布 $N(u_{z_t}, \sigma_{z_t}^2 I)$ 随机抽取到一个对角高斯向量 z_t ，将该对角高斯向量 z_t 作为该数值的第一辅助变量。

子步骤 S1023、根据预置解码程序对所述第一辅助变量进行解码，得到所述数值对应的第二高斯分布。

预置模型包括解码程序，在得到该数值的第一辅助变量时，通过预置解码程序对该第一辅助变量进行解码，得到该第一辅助变量解码后第二高斯分布。

在一实施例中，所述解码程序包括第二预置神经网络模型和第二预置全连接层，所述根据预置解码程序对所述第一辅助变量进行解码，得到所述数值对应的第二高斯分布，包括：根据所述第一辅助变量和第二预置神经网络模型，得到所述第二预置神经网络模型的第二隐藏变量；根据第二预置全连接层和所述第二隐藏变量，得到所述数值对应的第二高斯分布。

示范性的，如图 3 所示，该解码程序包括第二预置神经网络 GPU 和第二全连接层 h_2 ，在获取到该数值的第一辅助变量时，获取该第二预置神经网络 GPU 之前的第一隐藏变量，通过将该第一辅助变量和第一隐藏变量输入该第二预置神经网络，得到第二预置神经网络输出对应的第二隐藏变量。例如，获取到第一辅助变量 z_t ，获取第二预置神经网络 GPU 之前产生的第一隐藏变量 d_{t-1} ，将该 z_t 和 d_{t-1} 输入第二预置神经网络 GPU，通过该第一预置神经网络 GPU 中第一预置公式 $d_t = (1 - c_t^e) \times \tanh(w^e z_t + u^e (r_t^e \times d_{t-1}) + b^e) + c_t^e \times d_{t-1}$ ，其中， w^e 、

u^e 、 b^e 为为变量参数，通过获取第一神经网络或预置模型的参数矩阵，该参数矩阵是以 0 为中心分布的，随机抽取该参数矩阵中的参数。 c_t^e 是 GRU 中的更新门，用于决定如何将新输入的自变量与之前的时序信息结合起来。 r_t^e 是 GRU 中的重置门，用于决定保留多少此前的时序信息。通过第二预置公式得到 $r_t^e = \text{sigmoid}(w^{r^e} z_t + u^{r^e} d_{t-1} + b^{r^e})$ ，其中， w^{r^e} 、 u^{r^e} 、 b^{r^e} 为为变量参数，通过获取第一神经网络或预置模型的参数矩阵，该参数矩阵是以 0 为中心分布的，随机抽取该参数矩阵中的参数作为 w^{r^e} 、 u^{r^e} 、 b^{r^e} 的初始参数。通过第三预置公式得到 $c_t^e = \text{sigmoid}(w^{c^e} z_t + u^{c^e} d_{t-1} + b^{c^e})$ ，其中， w^{c^e} 、 u^{c^e} 、 b^{c^e} 为为变量参数，通过获取第一神经网络或预置模型的参数矩阵，该参数矩阵是以 0 为中心分布的，随机抽取该参数矩阵中的参数作为 w^{c^e} 、 u^{c^e} 、 b^{c^e} 的初始参数。

在得到该第二预置神经网络的第二隐藏变量时，将该第二隐藏变量输入第二全连接层，通过第二全连接层中的网络层以预置 ReLU 作为激活函数，分别经过下面两个线性变换和 *softplus* 变换得到均值和标准差。例如，通过第一全连接层的第一线性变换函数 $u_{x_t} = w^{u_z} h_1(d_t) + b^{u_z}$ ，得到该数值的第二均值。其中， $w^{u_z} h_1$ 、 b^{u_z} 为常数。通过第一全连接层的第二线性变换函数 $\sigma_{x_t} = \text{softplus}[w^{u_z} h_1(d_t) + b^{u_z}] + \varepsilon^{\sigma_z}$ ，得到该数值的第二标准差。其中， $w^{u_z} h_1$ 、 b^{u_z} 、 ε^{σ_z} 为常数， ε 是一个很小的常数，为了防止计算时数值溢出而设置。

在得到该数值的第二均值和第二标准差时，通过第二均值和第二标准差构建该数值的第二高斯分布。例如，在获取第二均值 u_{x_t} 和第一标准差 σ_{x_t} 时，构建该数值对应时刻的特定高斯分布 $N(u_{x_t}, \sigma_{x_t}^2 I)$ ，其中特定高斯分布 $N(u_{x_t}, \sigma_{x_t}^2 I)$ 为第二高斯分布。

子步骤 S1024、根据所述第一高斯分布和第二高斯分布训练所述预置模型的损失函数，生成对应的确定性模型。

通过得到该数值的第一高斯分布和第二高斯分布训练预置模型的损失函数，得到对应的确定性模型。例如，通过得到的第一高斯分布训练编码程序的第一损失函数，通过得到的第二高斯分布训练解码函数的第二损失函数，得到对应的确定性函数。

在一实施例中，所述根据所述第一高斯分布和第二高斯分布训练所述预置模型的损失函数，生成对应的确定性模型，包括：获取所述第二高斯分布中的累积分布概率；根据所述累积分布概率和第一预置损失函数，得到所述数值的重建概率；根据所述第一高斯分布，得到所述第一高斯分布的正则项；根据所述正则项和第二预置损失函数，得到所述数值的正则化项；基于所述重建概率和所述正则化项训练所述解码程序的网路参数和所述编码程序的网路参数，生成对应的确定性模型。

示范性的，重建概率通过计算原始数据 x_t 在重建分布 $x_t | z_t \sim N(u_{x_t}, \sigma_{x_t}^2 I)$ 中的累积分布概率，即 $L_{recon} = \sum_{i=1}^N \log[p(x_i | z_i)] = \sum_{i=1}^N \frac{(x_i - u_{x_i})^2}{2\sigma_{x_i}^2 I} \times \log(\sqrt{2\pi\sigma_{x_i}^2})$ 接下来是计算正则化项。由于编码过程中不是直接对辅助变量进行拟合，而是通过拟合辅助变量的均值和标准差，进而重采样得到 z_t ，因此整个重构过程受到噪声的影响。噪声强度，也就是拟合的标准差 σ_{z_t} 是通过神经网络计算得到，所以整个训练过程为了能更好的重建原始数据，会尽量压缩其标准差接近于 0。一旦标准差接近 0 采样的随机性消失，只会得到一个固定的样本，即 $z_t = u_{z_t}$ 。这种情况下，模型等同于一个普通的自编码器。

例如， $z_t | x_t$ 服从标准的正态分布，就避免了噪声（即 σ_{z_t} ）被压缩为零的问题，因此保证模型可以生成与原始数据不同的新样本。

因此，我们以多元独立正态分布与标准正态分布的 KL 散度 $KL[N(u, \sigma^2) \| N(0, I)]$ 作为

这个正则项加入损失函数中。 $L_{regul} = \sum_{i=1}^N KL[N(u_{z_i}, \sigma_{z_i}^2 I) \| N(0, I)] = \frac{1}{2} \sum_{i=1}^N (-\log \sigma_{z_i}^2 + u_{z_i}^2 + \sigma_{z_i}^2 - 1)$,

其中, N 是 z_i 的个数。

将上述两个损失函数进行合并, 得到总损失函数:

$$L = L_{recon} + L_{regul}, L = \sum_{i=1}^N \frac{(x_i - u_{x_i})^2}{2\sigma_{x_i}^2 I} \times \log(\sqrt{2\pi\sigma_{x_i}^2}) + \frac{1}{2} \sum_{i=1}^N (-\log \sigma_{z_i}^2 + u_{z_i}^2 + \sigma_{z_i}^2 - 1),$$

通过最大化上述损失函数来不断训练编码器及解码器的网络参数, 直至达到预期精度。

在一实施例中, 所述得到所述数值的重建概率之后, 还包括: 记录所述待训练的多元 KPI 时间序列中各个数值的重建概率, 并将各个所述数值的重建概率进行排序; 确定排序后所述重建概率中预置序号对应的目标重建概率, 并将所述目标重建概率作为预置阈值。

示范性的, 记录预置模型训练中最后一次计算 L_{recon} 的数据, 我们得到了每个原始数据对应的重建概率, 记作 $S_i = \log[p(x_i|z_i)]$, 其中, $i=1$ 至 N , 合并为一个包含 N 个元素的集合 $S = \{S_1, S_2, \dots, S_N\}$ 。重建概率 S_i 越高意味着输入 x_i 服从时间序列的正常模式, 因此可以高度可信地对其进行重构。概率越小, 说明能重构原始数据的可能性越小, 因此异常的可能性就越大。由于异常数值占比很小, 根据实际经验, 我们选取所有 S_i 数据的下 5% 分位数作为阈值。

步骤 S103、获取待检测的多元 KPI 时间序列。

示范性的, 通过预置切分策略获取待检测多元 KPI 时间序列。例如, 检测到多元 KPI 时间序列上出现新的数值时, 基于该预置切分策略中的目标窗口信息对该多元 KPI 时间序列进行切分, 获取包含出现新的数值的待检测多元 KPI 时间序列。

步骤 S104、根据所述待检测的多元 KPI 时间序列和所述确定性模型, 获取所述待检测的多元 KPI 时间序列中待检测数值的重建概率值

示范性的, 将该待检测多元 KPI 时间序列输入该确定性模型的, 获取待检测多元 KPI 时间序列中的待检测数值, 该待检测数值为出现新的数值。将该待检测数值输入该确定性模型的第一预置神经网络, 得到该第一预置神经网络输出的隐藏变量。将该隐藏变量输入第一预置全连接层, 获取第一预置全连接层变换得到的第一均值和第一标准差。根据第一均值和第一标准差, 得到该待检测数值的第一高斯分布。基于该第一高斯分布, 得到该待检测数值的辅助变量。基于第二预置神经网络和辅助变量, 得到该第二预置神经网络输出的隐藏变量。基于该第二预置神经网络输出的隐藏变量和第二预置全连接层, 得到第二预置全连接层变换得到的第二均值和第二标准差。基于的第二均值和第二标准差, 构建该待检测数值的第二高斯分布图, 得到该待检测数值的重建概率。

步骤 S105、根据预置阈值和所述待检测数值的重建概率值, 确定所述待检测数值对应的时刻是否为异常点。

将该待检测数值的重建概率值与预置阈值进行比对, 确定该待检测数值的时刻是否为异常点; 若该待检测数值的重建概率值小于该预置阈值, 则确定该待检测数值的时刻为异常点。

在本申请实施例中, 通过无标签的多元 KPI 时间序列训练预置模型, 生成对应的确定性模型, 通过该确定性模型对该检测的多元 KPI 时间序列进行检测, 确定该多元 KPI 时间序列中待检测数值对应的时刻是否为异常点。利用贝叶斯网络的知识根据不同的输入数据建立不同的高斯分布, 然后从中抽样获取对应潜变量, 再根据潜变量建立特定分布, 再次抽样得到重建数据, 因此可以生成与原始数据不同的变量, 增强了模型的鲁棒性; 在编码过程中, 融合了基于确定性的模型和基于随机性的模型, 充分捕获了潜在空间中辅助变量之间的时间依赖性, 因此可以更好地表示输入数据的分布, 更适用于多元 KPI 时序数据的

异常检测；扩展了具有随机变量的递归神经网络的建模功能，融合了 GRU 模型在充分捕捉时间序列数据的时间依赖性，提升了模型精度的同时大大降低模型复杂度，节约了运算过程中的 CPU 占用率，也减少了所需的存储空间。

本申请通过根据预置切分策略和待训练的多元 KPI 时间序列，确定对应的待训练的多段子多元 KPI 时间序列，其中，所述待训练的子多元 KPI 时间序列为无标签；根据所述待训练的子多元 KPI 时间序列训练预置模型，生成对应的确定性模型；获取待检测的多元 KPI 时间序列；根据所述待检测的多元 KPI 时间序列和所述确定性模型，获取所述待检测的多元 KPI 时间序列中待检测数值的重建概率值；根据预置阈值和所述待检测数值的重建概率值，确定所述待检测数值对应的时刻是否为异常点，实现通过无标签的多元 KPI 时间序列生成确定性模型，增强了模型的鲁棒性，使得该确定性模型更适用于多元 KPI 时序数据的异常检测，并提高了检测的准确率。

请参照图 4，图 4 为本申请实施例提供的一种多元 KPI 时间序列的检测装置的示意性框图。

如图 4 所示，该多元 KPI 时间序列的检测装置 400，包括：第一确定模块 401、生成模块 402、第一获取模块 403、第二获取模块 404、第二确定模块 405。

第一确定模块 401，用于根据预置切分策略和待训练的多元 KPI 时间序列，得到对应的待训练的多段子多元 KPI 时间序列，其中，所述待训练的子多元 KPI 时间序列为无标签；

生成模块 402，用于根据所述待训练的子多元 KPI 时间序列训练预置模型，生成对应的确定性模型；

第一获取模块 403，用于获取待检测的多元 KPI 时间序列；

第二获取模块 404，用于根据所述待检测的多元 KPI 时间序列和所述确定性模型，获取所述待检测的多元 KPI 时间序列中待检测数值的重建概率值；

第二确定模块 405，用于根据预置阈值和所述待检测数值的重建概率值，确定所述待检测数值对应的时刻是否为异常点。

其中，生成模块 402 具体还用于：

将所述待训练的子多元 KPI 时间序列输入所述预置模型，提取所述待训练的子多元 KPI 时间序列中的数值；

根据预置编码程序对所述数值进行编码，获取所述数值的第一高斯分布和第一辅助变量；

根据预置解码程序对所述第一辅助变量进行解码，得到所述数值对应的第二高斯分布；

根据所述第一高斯分布和第二高斯分布训练所述预置模型的损失函数，生成对应的确定性模型。

其中，生成模块 402 具体还用于：

根据所述数值和第一预置神经网络模型，得到所述第一预置神经网络模型的第一隐藏变量；

根据第一预置全连接层和所述第一隐藏变量，得到所述数值对应的第一高斯分布；

基于所述第一高斯分布得到所述数值的第一辅助变量。

其中，生成模块 402 具体还用于：

根据所述第一辅助变量和第二预置神经网络模型，得到所述第二预置神经网络模型的第二隐藏变量；

根据第二预置全连接层和所述第二隐藏变量，得到所述数值对应的第二高斯分布。

其中，生成模块 402 具体还用于：

获取所述第二高斯分布中的累积分布概率；

根据所述累积分布概率和第一预置损失函数，得到所述数值的重建概率；

根据所述第一高斯分布，得到所述第一高斯分布的正则项；

根据所述正则项和第二预置损失函数，得到所述数值的正则化项；

基于所述重建概率和所述正则化项训练所述解码程序的网络参数和所述编码程序的网络参数，生成对应的确定性模型。

其中，多元 KPI 时间序列的检测装置还用于：

记录所述待训练的子多元 KPI 时间序列中各个数值的重建概率，并将各个所述数值的重建概率进行排序；

确定排序后所述重建概率中预置序号对应的目标重建概率，并将所述目标重建概率作为预置阈值。

其中，第二确定模块 405 还用于：

将所述待检测数值的重建概率值与预置阈值进行比对，确定所述待检测数值的时刻是否为异常点；

若所述待检测数值的重建概率值小于所述预置阈值，则确定所述待检测数值的时刻为异常点。

需要说明的是，所属领域的技术人员可以清楚地了解到，为了描述的方便和简洁，上述描述的装置和各模块及单元的具体工作过程，可以参考前述多元 KPI 时间序列的检测方法实施例中的对应过程，在此不再赘述。

上述实施例提供的装置可以实现为一种计算机程序的形式，该计算机程序可以在如图 5 所示的计算机设备上运行。

请参阅图 5，图 5 为本申请实施例提供的一种计算机设备的结构示意图。该计算机设备可以为终端。

如图 5 所示，该计算机设备包括通过系统总线连接的处理器、存储器和网络接口，其中，存储器可以是易失性的，也可以是非易失性的。

非易失性存储介质可存储操作系统和计算机程序。该计算机程序包括程序指令，该程序指令被执行时，可使得处理器执行任意一种多元 KPI 时间序列的检测方法。

处理器用于提供计算和控制能力，支撑整个计算机设备的运行。

内存储器为非易失性存储介质中的计算机程序的运行提供环境，该计算机程序被处理器执行时，可使得处理器执行任意一种多元 KPI 时间序列的检测方法。

该网络接口用于进行网络通信，如发送分配的任务等。本领域技术人员可以理解，图 5 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计算机设备的限定，具体的计算机设备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。

应当理解的是，处理器可以是中央处理单元 (Central Processing Unit, CPU)，该处理器还可以是其他通用处理器、数字信号处理器 (Digital Signal Processor, DSP)、专用集成电路 (Application Specific Integrated Circuit, ASIC)、现场可编程门阵列 (Field-Programmable Gate Array, FPGA) 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。其中，通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

其中，在一个实施例中，所述处理器用于运行存储在存储器中的计算机程序，以实现如下步骤：

根据预置切分策略和待训练的多元 KPI 时间序列，确定对应的待训练的多段子多元 KPI 时间序列，其中，所述待训练的子多元 KPI 时间序列为无标签；

根据所述待训练的子多元 KPI 时间序列训练预置模型，生成对应的确定性模型；

获取待检测的多元 KPI 时间序列；

根据所述待检测的多元 KPI 时间序列和所述确定性模型，获取所述待检测的多元 KPI 时间序列中待检测数值的重建概率值；

根据预置阈值和所述待检测数值的重建概率值，确定所述待检测数值对应的时刻是否为异常点。

在一个实施例中，所述处理器根据所述待训练的子多元 KPI 时间序列训练预置模型，生成对应的确定性模型实现时，用于实现：

将所述待训练的子多元 KPI 时间序列输入所述预置模型，提取所述待训练的子多元 KPI 时间序列中的数值；

根据预置编码程序对所述数值进行编码，获取所述数值的第一高斯分布和第一辅助变量；

根据预置解码程序对所述第一辅助变量进行解码，得到所述数值对应的第二高斯分布；

根据所述第一高斯分布和第二高斯分布训练所述预置模型的损失函数，生成对应的确定性模型。

在一个实施例中，所述处理器所述预置编码程序包括第一预置神经网络模型和第一全连接层，所述根据预置编码程序对所述数值进行编码，获取所述数值的第一高斯分布和第一辅助变量实现时，用于实现：

根据所述数值和第一预置神经网络模型，得到所述第一预置神经网络模型的第一隐藏变量；

根据第一预置全连接层和所述第一隐藏变量，得到所述数值对应的第一高斯分布；

基于所述第一高斯分布得到所述数值的第一辅助变量。

在一个实施例中，所述处理器所述解码程序包括第二预置神经网络模型和第二预置全连接层，所述根据预置解码程序对所述第一辅助变量进行解码，得到所述数值对应的第二高斯分布实现时，用于实现：

根据所述第一辅助变量和第二预置神经网络模型，得到所述第二预置神经网络模型的第二隐藏变量；

根据第二预置全连接层和所述第二隐藏变量，得到所述数值对应的第二高斯分布。

在一个实施例中，所述处理器所述根据所述第一高斯分布和第二高斯分布训练所述预置模型的损失函数，生成对应的确定性模型实现时，用于实现：

获取所述第二高斯分布中的累积分布概率；

根据所述累积分布概率和第一预置损失函数，得到所述数值的重建概率；

根据所述第一高斯分布，得到所述第一高斯分布的正则项；

根据所述正则项和第二预置损失函数，得到所述数值的正则化项；

基于所述重建概率和所述正则化项训练所述解码程序的网络参数和所述编码程序的网络参数，生成对应的确定性模型。

在一个实施例中，所述处理器所述得到所述数值的重建概率之后实现时，用于实现：

记录所述待训练的子多元 KPI 时间序列中各个数值的重建概率，并将各个所述数值的重建概率进行排序；

确定排序后所述重建概率中预置序号对应的目标重建概率，并将所述目标重建概率作为预置阈值。

在一个实施例中，所述处理器所述根据预置阈值和所述待检测数值的重建概率值，确定所述待检测数值的时刻是否为异常点实现时，用于实现：

将所述待检测数值的重建概率值与预置阈值进行比对，确定所述待检测数值的时刻是否为异常点；

若所述待检测数值的重建概率值小于所述预置阈值，则确定所述待检测数值的时刻为

异常点。

本申请实施例还提供一种计算机可读存储介质，计算机可读存储介质可以是易失性的，也可以是非易失性的，所述计算机可读存储介质上存储有计算机程序，所述计算机程序中包括程序指令，所述程序指令被执行时所实现的方法可参照本申请多元 KPI 时间序列的检测方法的各个实施例。

其中，所述计算机可读存储介质可以是前述实施例所述的计算机设备的内部存储单元，例如所述计算机设备的硬盘或内存。所述计算机可读存储介质也可以是所述计算机设备的外部存储设备，例如所述计算机设备上配备的插接式硬盘，智能存储卡(Smart Media Card, SMC)，安全数字(Secure Digital, SD)卡，闪存卡(Flash Card)等。

进一步地，所述计算机可读存储介质可主要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等；存储数据区可存储根据区块链节点的使用所创建的数据等。

本申请所指区块链是预置模型和确定性模型的存储、点对点传输、共识机制、加密算法等计算机技术的新型应用模式。区块链(Blockchain)，本质上是一个去中心化的数据库，是一串使用密码学方法相关联产生的数据块，每一个数据块中包含了一批次网络交易的信息，用于验证其信息的有效性(防伪)和生成下一个区块。区块链可以包括区块链底层平台、平台产品服务层以及应用服务层等。

需要说明的是，在本文中，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过程、方法、物品或者系统不仅包括那些要素，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过程、方法、物品或者系统所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括该要素的过程、方法、物品或者系统中还存在另外的相同要素。

上述本申请实施例序号仅仅为了描述，不代表实施例的优劣。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，任何熟悉本技术领域的技术人员在本申请揭露的技术范围内，可轻易想到各种等效的修改或替换，这些修改或替换都应涵盖在本申请的保护范围之内。因此，本申请的保护范围应以权利要求的保护范围为准。

权利要求书

1、一种多元 KPI 时间序列的检测方法，其中，包括：

根据预置切分策略和待训练的多元 KPI 时间序列，确定对应的待训练的多段子多元 KPI 时间序列，其中，所述待训练的子多元 KPI 时间序列为无标签；

根据所述待训练的子多元 KPI 时间序列训练预置模型，生成对应的确定性模型；

获取待检测的多元 KPI 时间序列；

根据所述待检测的多元 KPI 时间序列和所述确定性模型，获取所述待检测的多元 KPI 时间序列中待检测数值的重建概率值；

根据预置阈值和所述待检测数值的重建概率值，确定所述待检测数值对应的时刻是否为异常点。

2、如权利要求 1 所述的多元 KPI 时间序列的检测方法，其中，所述根据所述待训练的子多元 KPI 时间序列训练预置模型，生成对应的确定性模型，包括：

将所述待训练的子多元 KPI 时间序列输入所述预置模型，提取所述待训练的子多元 KPI 时间序列中的数值；

根据预置编码程序对所述数值进行编码，获取所述数值的第一高斯分布和第一辅助变量；

根据预置解码程序对所述第一辅助变量进行解码，得到所述数值对应的第二高斯分布；

根据所述第一高斯分布和第二高斯分布训练所述预置模型的损失函数，生成对应的确定性模型。

3、如权利要求 2 所述的多元 KPI 时间序列的检测方法，其中，所述预置编码程序包括第一预置神经网络模型和第一全连接层，所述根据预置编码程序对所述数值进行编码，获取所述数值的第一高斯分布和第一辅助变量，包括：

根据所述数值和第一预置神经网络模型，得到所述第一预置神经网络模型的第一隐藏变量；

根据第一预置全连接层和所述第一隐藏变量，得到所述数值对应的第一高斯分布；

基于所述第一高斯分布得到所述数值的第一辅助变量。

4、如权利要求 2 所述的多元 KPI 时间序列的检测方法，其中，所述解码程序包括第二预置神经网络模型和第二预置全连接层，所述根据预置解码程序对所述第一辅助变量进行解码，得到所述数值对应的第二高斯分布，包括：

根据所述第一辅助变量和第二预置神经网络模型，得到所述第二预置神经网络模型的第二隐藏变量；

根据第二预置全连接层和所述第二隐藏变量，得到所述数值对应的第二高斯分布。

5、如权利要求 2 所述的多元 KPI 时间序列的检测方法，其中，所述根据所述第一高斯分布和第二高斯分布训练所述预置模型的损失函数，生成对应的确定性模型，包括：

获取所述第二高斯分布中的累积分布概率；

根据所述累积分布概率和第一预置损失函数，得到所述数值的重建概率；

根据所述第一高斯分布，得到所述第一高斯分布的正则项；

根据所述正则项和第二预置损失函数，得到所述数值的正则化项；

基于所述重建概率和所述正则化项训练所述解码程序的神经网络参数和所述编码程序的神经网络参数，生成对应的确定性模型。

6、如权利要求 5 所述的多元 KPI 时间序列的检测方法，其中，所述得到所述数值的重建概率之后，还包括：

记录所述待训练的子多元 KPI 时间序列中各个数值的重建概率，并将各个所述数值的重建概率进行排序；

确定排序后所述重建概率中预置序号对应的目标重建概率，并将所述目标重建概率作为预置阈值。

7、如权利要求 1 所述的多元 KPI 时间序列的检测方法，其中，所述根据预置阈值和所述待检测数值的重建概率值，确定所述待检测数值的时刻是否为异常点，包括：

将所述待检测数值的重建概率值与预置阈值进行比对，确定所述待检测数值的时刻是否为异常点；

若所述待检测数值的重建概率值小于所述预置阈值，则确定所述待检测数值的时刻为异常点。

8、一种多元 KPI 时间序列的检测装置，其中，包括：

第一确定模块，用于根据预置切分策略和待训练的多元 KPI 时间序列，确定对应的待训练的多段子多元 KPI 时间序列，其中，所述待训练的子多元 KPI 时间序列为无标签；

生成模块，用于根据所述待训练的子多元 KPI 时间序列训练预置模型，生成对应的确定性模型；

第一获取模块，用于获取待检测的多元 KPI 时间序列；

第二获取模块，用于根据所述待检测的多元 KPI 时间序列和所述确定性模型，获取所述待检测的多元 KPI 时间序列中待检测数值的重建概率值；

第二确定模块，用于根据预置阈值和所述待检测数值的重建概率值，确定所述待检测数值对应的时刻是否为异常点。

9、一种计算机设备，其中，所述计算机设备包括处理器、存储器、以及存储在所述存储器上并可被所述处理器执行的计算机程序，其中所述计算机程序被所述处理器执行时，实现如下步骤：

根据预置切分策略和待训练的多元 KPI 时间序列，确定对应的待训练的多段子多元 KPI 时间序列，其中，所述待训练的子多元 KPI 时间序列为无标签；

根据所述待训练的子多元 KPI 时间序列训练预置模型，生成对应的确定性模型；

获取待检测的多元 KPI 时间序列；

根据所述待检测的多元 KPI 时间序列和所述确定性模型，获取所述待检测的多元 KPI 时间序列中待检测数值的重建概率值；

根据预置阈值和所述待检测数值的重建概率值，确定所述待检测数值对应的时刻是否为异常点。

10、如权利要求 9 所述的计算机设备，其中，所述根据所述待训练的子多元 KPI 时间序列训练预置模型，生成对应的确定性模型，包括：

将所述待训练的子多元 KPI 时间序列输入所述预置模型，提取所述待训练的子多元 KPI 时间序列中的数值；

根据预置编码程序对所述数值进行编码，获取所述数值的第一高斯分布和第一辅助变量；

根据预置解码程序对所述第一辅助变量进行解码，得到所述数值对应的第二高斯分布；

根据所述第一高斯分布和第二高斯分布训练所述预置模型的损失函数，生成对应的确定性模型。

11、如权利要求 10 所述的计算机设备，其中，所述预置编码程序包括第一预置神经网络模型和第一全连接层，所述根据预置编码程序对所述数值进行编码，获取所述数值的第一高斯分布和第一辅助变量，包括：

根据所述数值和第一预置神经网络模型，得到所述第一预置神经网络模型的第一隐藏变量；

根据第一预置全连接层和所述第一隐藏变量，得到所述数值对应的第一高斯分布；

基于所述第一高斯分布得到所述数值的第一辅助变量。

12、如权利要求 10 所述的计算机设备，其中，所述解码程序包括第二预置神经网络模型和第二预置全连接层，所述根据预置解码程序对所述第一辅助变量进行解码，得到所述数值对应的第二高斯分布，包括：

根据所述第一辅助变量和第二预置神经网络模型，得到所述第二预置神经网络模型的第二隐藏变量；

根据第二预置全连接层和所述第二隐藏变量，得到所述数值对应的第二高斯分布。

13、如权利要求 10 所述的计算机设备，其中，所述根据所述第一高斯分布和第二高斯分布训练所述预置模型的损失函数，生成对应的确定性模型，包括：

获取所述第二高斯分布中的累积分布概率；

根据所述累积分布概率和第一预置损失函数，得到所述数值的重建概率；

根据所述第一高斯分布，得到所述第一高斯分布的正则项；

根据所述正则项和第二预置损失函数，得到所述数值的正则化项；

基于所述重建概率和所述正则化项训练所述解码程序的网络参数和所述编码程序的网络参数，生成对应的确定性模型。

14、如权利要求 13 所述的计算机设备，其中，所述得到所述数值的重建概率之后，还实现以下步骤：

记录所述待训练的子多元 KPI 时间序列中各个数值的重建概率，并将各个所述数值的重建概率进行排序；

确定排序后所述重建概率中预置序号对应的目标重建概率，并将所述目标重建概率作为预置阈值。

15、如权利要求 9 所述的计算机设备，其中，所述根据预置阈值和所述待检测数值的重建概率值，确定所述待检测数值的时刻是否为异常点，包括：

将所述待检测数值的重建概率值与预置阈值进行比对，确定所述待检测数值的时刻是否为异常点；

若所述待检测数值的重建概率值小于所述预置阈值，则确定所述待检测数值的时刻为异常点。

16、一种计算机可读存储介质，其中，所述计算机可读存储介质上存储有计算机程序，其中所述计算机程序被处理器执行时，实现如下步骤：

根据预置切分策略和待训练的多元 KPI 时间序列，确定对应的待训练的多段子多元 KPI 时间序列，其中，所述待训练的子多元 KPI 时间序列为无标签；

根据所述待训练的子多元 KPI 时间序列训练预置模型，生成对应的确定性模型；

获取待检测的多元 KPI 时间序列；

根据所述待检测的多元 KPI 时间序列和所述确定性模型，获取所述待检测的多元 KPI 时间序列中待检测数值的重建概率值；

根据预置阈值和所述待检测数值的重建概率值，确定所述待检测数值对应的时刻是否为异常点。

17、如权利要求 16 所述的计算机可读存储介质，其中，所述根据所述待训练的子多元 KPI 时间序列训练预置模型，生成对应的确定性模型，包括：

将所述待训练的子多元 KPI 时间序列输入所述预置模型，提取所述待训练的子多元 KPI 时间序列中的数值；

根据预置编码程序对所述数值进行编码，获取所述数值的第一高斯分布和第一辅助变量；

根据预置解码程序对所述第一辅助变量进行解码，得到所述数值对应的第二高斯分布；

根据所述第一高斯分布和第二高斯分布训练所述预置模型的损失函数，生成对应的确定性模型。

18、如权利要求 17 所述的计算机可读存储介质，其中，所述预置编码程序包括第一预置神经网络模型和第一全连接层，所述根据预置编码程序对所述数值进行编码，获取所述数值的第一高斯分布和第一辅助变量，包括：

根据所述数值和第一预置神经网络模型，得到所述第一预置神经网络模型的第一隐藏变量；

根据第一预置全连接层和所述第一隐藏变量，得到所述数值对应的第一高斯分布；

基于所述第一高斯分布得到所述数值的第一辅助变量。

19、如权利要求 17 所述的计算机可读存储介质，其中，所述解码程序包括第二预置神经网络模型和第二预置全连接层，所述根据预置解码程序对所述第一辅助变量进行解码，得到所述数值对应的第二高斯分布，包括：

根据所述第一辅助变量和第二预置神经网络模型，得到所述第二预置神经网络模型的第二隐藏变量；

根据第二预置全连接层和所述第二隐藏变量，得到所述数值对应的第二高斯分布。

20、如权利要求 17 所述的计算机可读存储介质，其中，所述根据所述第一高斯分布和第二高斯分布训练所述预置模型的损失函数，生成对应的确定性模型，包括：

获取所述第二高斯分布中的累积分布概率；

根据所述累积分布概率和第一预置损失函数，得到所述数值的重建概率；

根据所述第一高斯分布，得到所述第一高斯分布的正则项；

根据所述正则项和第二预置损失函数，得到所述数值的正则化项；

基于所述重建概率和所述正则化项训练所述解码程序的网络参数和所述编码程序的网络参数，生成对应的确定性模型。

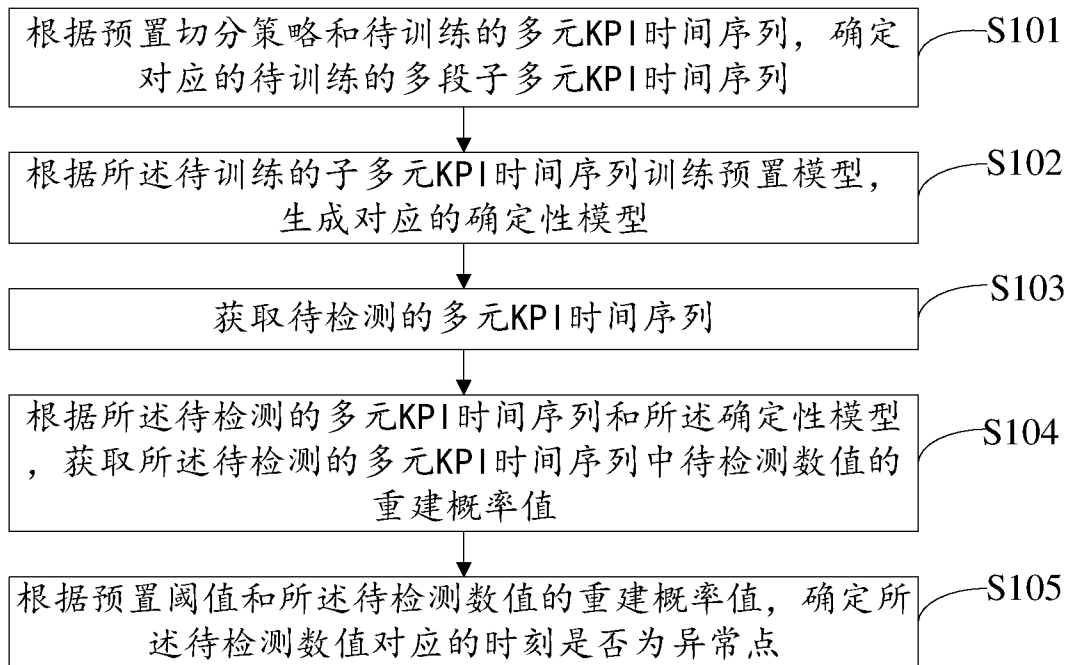


图 1

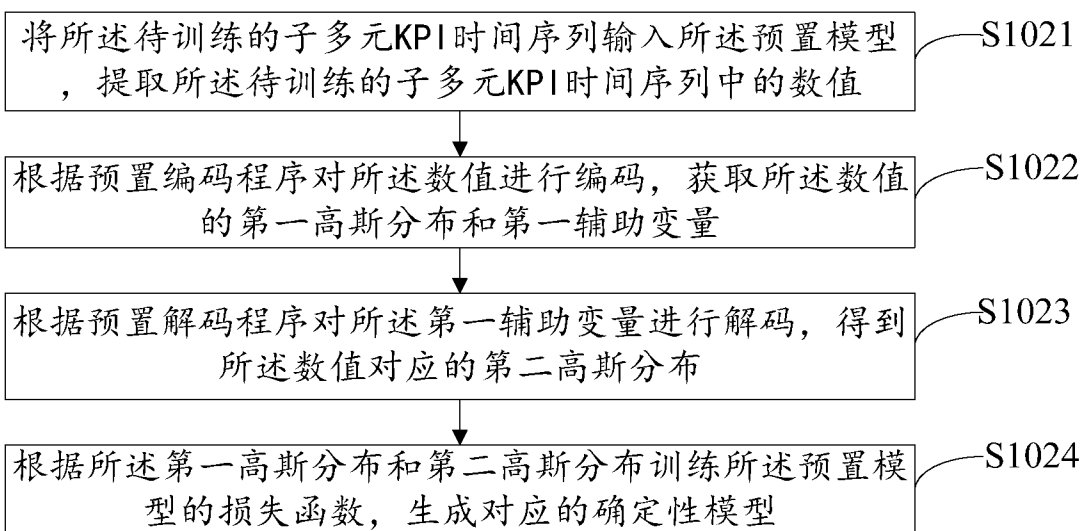


图 2

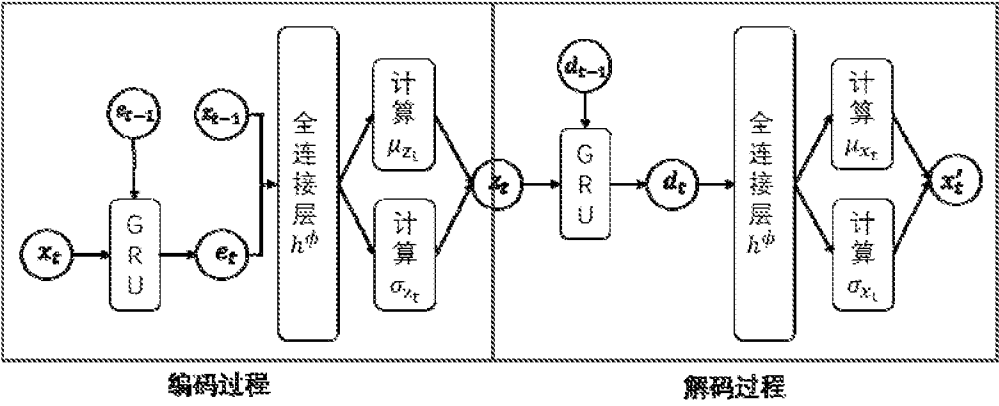


图 3

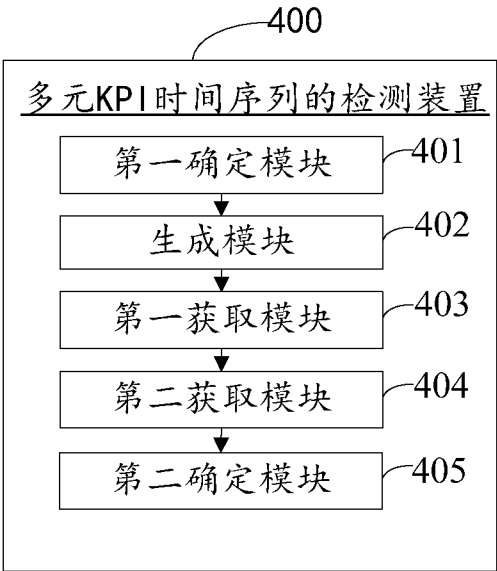


图 4

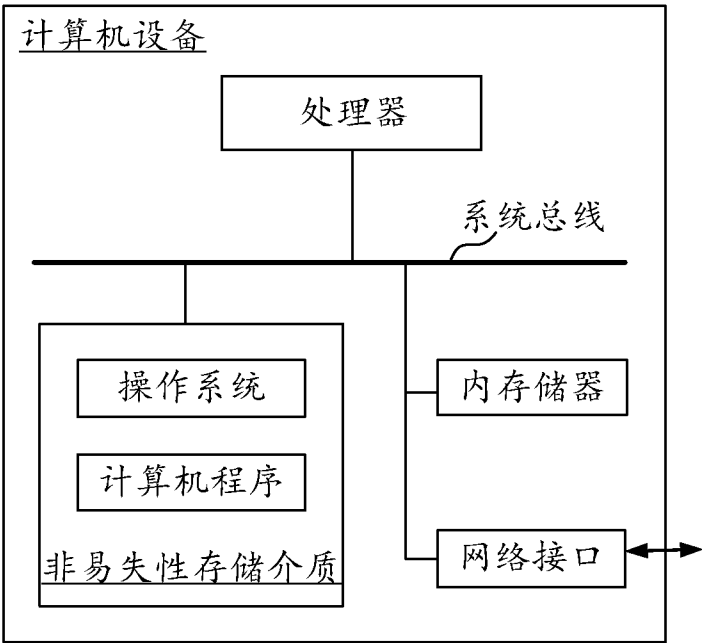


图 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/125003

A. CLASSIFICATION OF SUBJECT MATTER

G06F 16/2458(2019.01)i; G06N 3/04(2006.01)n

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, G06N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI, EPODOC, CNPAT, CNKI, BAIDU, IEEE: 多元, 时间序列, 异常, 无标签, 模型, 重建, 重构, 概率, 编码, 解码, 损失, 神经网络, 高斯, multivariate, time series, anomaly, unlabeled, model, reconstruction, probability, coding, decoding, loss, neural network, Gaussian

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	陈俊生 等 (CHEN, Junsheng et al.). "采用滑动窗口及多重加噪比堆栈降噪自编码的风电机组状态异常检测方法 (A Method for Detecting Anomaly Conditions of Wind Turbines Using Stacked Denoising Autoencoders with Sliding Window and Multiple Noise Ratios)" 电工技术学报 (Transactions of China Electrotechnical Society), Vol. 35, No. 2, 31 January 2020 (2020-01-31), abstract, and pp. 347-352	1, 7-9, 15, 16
A	CN 109492193 A (TONGJI UNIVERSITY) 19 March 2019 (2019-03-19) entire document	1-20
A	CN 111277603 A (HANGZHOU DPTECH TECHNOLOGIES CO., LTD.) 12 June 2020 (2020-06-12) entire document	1-20
A	CN 109543943 A (ELECTRIC POWER RESEARCH INSTITUTION OF STATE GRID LIAONING ELECTRIC POWER SUPPLY CO., LTD. et al.) 29 March 2019 (2019-03-29) entire document	1-20
A	WO 2020147480 A1 (ALIBABA GROUP HOLDING LIMITED) 23 July 2020 (2020-07-23) 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

10 June 2021

Date of mailing of the international search report

18 June 2021

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/CN2020/125003

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	109492193	A	19 March 2019	CN	109492193	B	27 November 2020
CN	111277603	A	12 June 2020	None			
CN	109543943	A	29 March 2019	None			
WO	2020147480	A1	23 July 2020	CN	110058977	A	26 July 2019
				CN	110058977	B	14 August 2020

国际检索报告

国际申请号

PCT/CN2020/125003

A. 主题的分类 G06F 16/2458(2019.01)i; G06N 3/04(2006.01)n 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, G06N 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用)) WPI, EPDOC, CNPAT, CNKI, BAIDU, IEEE: 多元, 时间序列, 异常, 无标签, 模型, 重建, 重构, 概率, 编码, 解码, 损失, 神经网络, 高斯, multivariate, time series, anomaly, unlabeled, model, reconstruction, probability, coding, decoding, loss, neural network, Gaussian																				
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>陈俊生 等. “采用滑动窗口及多重加噪比堆栈降噪自编码的风电机组状态异常检测方法” 《电工技术学报》, 第35卷, 第2期, 2020年 1月 31日 (2020 - 01 - 31), 摘要, 第347-352页</td> <td>1, 7-9, 15, 16</td> </tr> <tr> <td>A</td> <td>CN 109492193 A (同济大学) 2019年 3月 19日 (2019 - 03 - 19) 全文</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 111277603 A (杭州迪普科技股份有限公司) 2020年 6月 12日 (2020 - 06 - 12) 全文</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 109543943 A (国网辽宁省电力有限公司电力科学研究院 等) 2019年 3月 29日 (2019 - 03 - 29) 全文</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>W0 2020147480 A1 (阿里巴巴集团控股有限公司) 2020年 7月 23日 (2020 - 07 - 23) 全文</td> <td>1-20</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	X	陈俊生 等. “采用滑动窗口及多重加噪比堆栈降噪自编码的风电机组状态异常检测方法” 《电工技术学报》, 第35卷, 第2期, 2020年 1月 31日 (2020 - 01 - 31), 摘要, 第347-352页	1, 7-9, 15, 16	A	CN 109492193 A (同济大学) 2019年 3月 19日 (2019 - 03 - 19) 全文	1-20	A	CN 111277603 A (杭州迪普科技股份有限公司) 2020年 6月 12日 (2020 - 06 - 12) 全文	1-20	A	CN 109543943 A (国网辽宁省电力有限公司电力科学研究院 等) 2019年 3月 29日 (2019 - 03 - 29) 全文	1-20	A	W0 2020147480 A1 (阿里巴巴集团控股有限公司) 2020年 7月 23日 (2020 - 07 - 23) 全文	1-20
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																		
X	陈俊生 等. “采用滑动窗口及多重加噪比堆栈降噪自编码的风电机组状态异常检测方法” 《电工技术学报》, 第35卷, 第2期, 2020年 1月 31日 (2020 - 01 - 31), 摘要, 第347-352页	1, 7-9, 15, 16																		
A	CN 109492193 A (同济大学) 2019年 3月 19日 (2019 - 03 - 19) 全文	1-20																		
A	CN 111277603 A (杭州迪普科技股份有限公司) 2020年 6月 12日 (2020 - 06 - 12) 全文	1-20																		
A	CN 109543943 A (国网辽宁省电力有限公司电力科学研究院 等) 2019年 3月 29日 (2019 - 03 - 29) 全文	1-20																		
A	W0 2020147480 A1 (阿里巴巴集团控股有限公司) 2020年 7月 23日 (2020 - 07 - 23) 全文	1-20																		
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																				
<table border="0"> <tr> <td> * 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 </td> <td> “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 </td> </tr> </table>			* 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																
* 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																		
2021年 6月 10日		2021年 6月 18日																		
ISA/CN的名称和邮寄地址		授权官员																		
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		吴雪 电话号码 86-(10)-53961406																		

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2020/125003

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	109492193	A	2019年 3月 19日	CN 109492193 B	2020年 11月 27日
CN	111277603	A	2020年 6月 12日	无	
CN	109543943	A	2019年 3月 29日	无	
WO	2020147480	A1	2020年 7月 23日	CN 110058977 A	2019年 7月 26日
				CN 110058977 B	2020年 8月 14日