



(12) 发明专利申请

(10) 申请公布号 CN 114064383 A

(43) 申请公布日 2022. 02. 18

(21) 申请号 202010761891.8

(22) 申请日 2020.07.31

(71) 申请人 中移(苏州)软件技术有限公司
地址 215163 江苏省苏州市高新区昆仑山路58号1幢

申请人 中移互联网有限公司
中国移动通信集团有限公司

(72) 发明人 胡洪涛 徐景成 胡建华 朱耀磷
邓宇辉

(74) 专利代理机构 北京派特恩知识产权代理有限公司 11270
代理人 姚文娴 张颖玲

(51) Int. Cl.
G06F 11/30 (2006.01)
G06N 3/08 (2006.01)

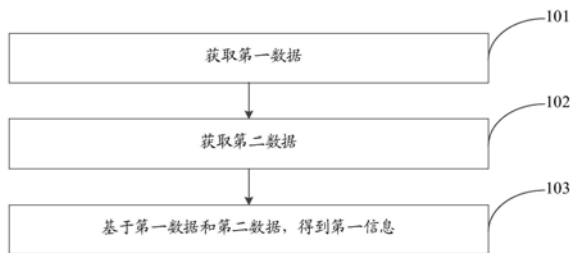
权利要求书2页 说明书19页 附图3页

(54) 发明名称

一种信息处理方法、系统、设备及计算机存储介质

(57) 摘要

本申请公开了一种信息处理方法:获取第一数据;其中,第一数据,包括数据库第一时刻的至少一种运行状态信息和第一时段的至少一个运行参数;第一时刻,用于表示数据库历史运行过程中的时刻;第一时段,表示第一时刻之前预设长度的时间段;获取第二数据;第二数据,用于表示数据库当前时刻的至少一个运行参数;基于第一数据和第二数据,得到第一信息;第一信息,用于表示数据库的第二时刻和/或第二时段的运行状态信息;第二时刻,用于表示数据库运行的未来时刻;第二时段,用于表示数据库运行的预设长度的未来时段。以上方法可以预测数据库未来时刻和/或未来时段的运行状态。本申请还公开了一种信息处理系统、设备和计算机存储介质。



1. 一种信息处理方法,其特征在于,所述方法包括:

获取第一数据;其中,所述第一数据,包括数据库第一时刻的至少一种运行状态信息和第一时段的所述至少一个运行参数;其中,所述第一时刻,用于表示所述数据库历史运行过程中的时刻;所述第一时段,用于表示所述第一时刻之前预设长度的时间段;

获取第二数据;其中,所述第二数据,用于表示所述数据库当前时刻的所述至少一个运行参数;

基于所述第一数据和所述第二数据,得到第一信息;其中,所述第一信息,用于表示所述数据库的第二时刻和/或第二时段的运行状态信息;所述第二时刻,用于表示所述数据库运行的未来时刻;所述第二时段,用于表示所述数据库运行的预设长度的未来时段。

2. 根据权利要求1所述的方法,其特征在于,所述基于所述第一数据和所述第二数据,得到所述第一信息,包括:

获取信息分类规则;

基于所述信息分类规则,对所述第一数据中的所述至少一个运行参数进行处理,得到至少一个第一序列;其中,所述第一序列,用于表示在所述第一时段所述至少一个运行参数的状态变化的序列;

对所述至少一种运行状态信息和所述至少一个第一序列进行组合,得到至少一个第二序列;

基于所述至少一个第二序列,对所述第二数据进行处理,得到所述第一信息。

3. 根据权利要求2所述的方法,其特征在于,所述基于所述信息分类规则,对所述第一数据中的所述至少一个运行参数进行处理,得到至少一个第一序列,包括:

基于所述信息分类规则,得到第一分类规则和第二分类规则;其中,所述第一分类规则,用于表示所述第一数据中的所述至少一个运行参数的类别划分规则;所述第二分类规则,用于表示所述第一数据中的所述至少一个运行参数的参数值的区间划分规则;

基于所述第一分类规则 and 所述第二分类规则,对所述第一数据中的所述运行参数进行处理,得到至少一个第一序列。

4. 根据权利要求2所述的方法,其特征在于,所述基于所述至少一个第二序列,对所述第二数据进行处理,得到所述第一信息,包括:

获取训练完成的循环神经网络;其中,所述循环神经网络,用于预测所述数据库的所述第二时刻和/或所述第二时段的所述至少一个运行状态参数值;

基于所述训练完成的循环神经网络和所述至少一个第二序列,对所述第二数据进行处理,得到所述第一信息。

5. 根据权利要求4所述的方法,其特征在于,所述获取训练完成的循环神经网络,包括:

基于所述第一数据,对所述循环神经网络进行训练,得到所述训练完成的循环神经网络。

6. 根据权利要求4所述的方法,其特征在于,所述基于所述训练完成的循环神经网络和所述至少一个第二序列,对所述第二数据进行处理,得到所述第一信息,包括:

基于所述训练完成的循环神经网络,对所述第二数据进行处理,得到至少一个第三序列;其中,所述至少一个第三序列,包括所述数据库的所述第二时刻和/或所述第二时段的所述至少一个运行参数值;

基于所述至少一个第三序列和所述至少一个第二序列,得到所述第一信息。

7. 根据权利要求6所述的方法,其特征在于,所述基于所述至少一个第三序列和所述至少一个第二序列,得到所述第一信息,包括:

对所述至少一个第二序列进行聚类,得到至少一个类簇;

基于所述至少一个类簇,对所述至少一个第三序列进行处理,得到所述第一信息。

8. 根据权利要求6所述的方法,其特征在于,所述基于所述训练完成的循环神经网络,对所述第二数据进行处理,得到至少一个第三序列,包括:

将所述第二数据输入至所述训练完成的循环神经网络,得到第三数据;

基于所述信息分类规则,对所述第三数据进行处理,得到至少一个第三序列。

9. 一种信息处理系统,其特征在于,所述系统包括:处理器、存储器和通信总线;其中,所述通信总线用于实现处理器和存储器之间的通信连接;

所述处理器用于执行存储器中的信息处理方法的程序,以实现以下步骤:

获取第一数据;其中,所述第一数据,包括数据库第一时刻的至少一种运行状态信息和第一时段的所述至少一个运行参数信息;其中,所述第一时刻,用于表示所述数据库历史运行过程中的时刻;所述第一时段,用于表示所述第一时刻之前预设长度的时间段;

获取第二数据;其中,所述第二数据,用于表示所述数据库当前时刻的所述至少一个运行参数信息;

基于所述第一数据和所述第二数据,得到第一信息;其中,所述第一信息,用于表示所述数据库第二时刻和/或第二时段的运行状态信息;其中,所述第二时刻,用于表示所述数据库运行的未来时刻;所述第二时段,用于表示所述数据库运行的预设长度的未来时段。

10. 一种信息处理设备,其特征在于,所述信息处理设备包括:获取模块和处理模块;其中:

所述获取模块,用于获取第一数据和第二数据;其中,所述第一数据,包括数据库第一时刻的至少一种运行状态信息和第一时段的所述至少一个运行参数;其中,所述第一时刻,用于表示所述数据库历史运行过程中的时刻;所述第一时段,用于表示所述第一时刻之前预设长度的时间段;所述第二数据,用于表示所述数据库当前时刻的所述至少一个运行参数;

所述处理模块,用于基于所述第一数据和所述第二数据,得到第一信息;其中,所述第一信息,用于表示所述数据库第二时刻和/或第二时段的运行状态信息;其中,所述第二时刻,用于表示所述数据库运行的未来时刻;所述第二时段,用于表示所述数据库运行的预设长度的未来时段。

11. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质存储有一个或者多个程序,所述一个或者多个程序可被一个或者多个处理器执行,以实现如权利要求1至8中任一项所述的信息处理方法的步骤。

一种信息处理方法、系统、设备及计算机存储介质

技术领域

[0001] 本申请涉及电子信息技术领域,尤其涉及一种信息处理方法、系统、设备以及计算机可读存储介质。

背景技术

[0002] 随着数据库的长期高负荷多并发运行,数据库的运行状态难免会出现异常。在相关技术中,通常通过监控表征数据库运行状态的关键数据,并根据这些关键数据来确定当前数据库的运行状态是否正常。然而,以上方案仅能根据当前实际检测到的数据库关键数据,来判断数据库当前的运行状态是否出现异常,而无法在数据库尚未出现异常的时候,对数据库的未来时刻或未来时段是否会出现异常进行预测。

发明内容

[0003] 本申请提供了一种信息处理方法,该方法可以实现数据库未来时刻运行状态信息的预测,从而解决了相关技术中在数据库尚未出现异常时候无法预测数据库在未来的运行中是否会出现异常的问题。

[0004] 本申请提供了一种信息处理方法,所述方法包括:

[0005] 获取第一数据;其中,所述第一数据,包括数据库第一时刻的至少一种运行状态信息和第一时段的所述至少一个运行参数;其中,所述第一时刻,用于表示所述数据库历史运行过程中的时刻;所述第一时段,用于表示所述第一时刻之前预设长度的时间段;

[0006] 获取第二数据;其中,所述第二数据,用于表示所述数据库当前时刻的所述至少一个运行参数;

[0007] 基于所述第一数据和所述第二数据,得到第一信息;其中,所述第一信息,用于表示所述数据库的第二时刻和/或第二时段的运行状态信息;所述第二时刻,用于表示所述数据库运行的未来时刻;所述第二时段,用于表示所述数据库运行的预设长度的未来时段。

[0008] 在一些实施方式中,所述基于所述第一数据和所述第二数据,得到所述第一信息,包括:

[0009] 获取信息分类规则;

[0010] 基于所述信息分类规则,对所述第一数据中的所述至少一个运行参数进行处理,得到至少一个第一序列;其中,所述第一序列,用于表示所述在第一时段所述至少一个运行参数的状态变化的序列;

[0011] 对所述至少一种运行状态信息和所述至少一个第一序列进行组合,得到至少一个第二序列;

[0012] 基于所述至少一个第二序列,对所述第二数据进行处理,得到所述第一信息。

[0013] 在一些实施方式中,所述基于所述信息分类规则,对所述第一数据中的所述至少一个运行参数进行处理,得到至少一个第一序列,包括:

[0014] 基于所述信息分类规则,得到第一分类规则和第二分类规则;其中,所述第一分类

规则,用于表示所述第一数据中的所述至少一个运行参数的类别划分规则;所述第二分类规则,用于表示所述第一数据中的所述至少一个运行参数的参数值的区间划分规则;

[0015] 基于所述第一分类规则和所述第二分类规则,对所述第一数据中的所述运行参数进行处理,得到至少一个第一序列。

[0016] 在一些实施方式中,所述基于所述至少一个第二序列,对所述第二数据进行处理,得到所述第一信息,包括:

[0017] 获取训练完成的循环神经网络;其中,所述循环神经网络,用于预测所述数据库的所述第二时刻和/或所述第二时段的所述至少一个运行状态参数值;

[0018] 基于所述训练完成的循环神经网络和所述至少一个第二序列,对所述第二数据进行处理,得到所述第一信息。

[0019] 在一些实施方式中,所述获取训练完成的循环神经网络,包括:

[0020] 基于所述第一数据,对所述循环神经网络进行训练,得到所述训练完成的循环神经网络。

[0021] 在一些实施方式中,所述基于所述训练完成的循环神经网络和所述至少一个第二序列,对所述第二数据进行处理,得到所述第一信息,包括:

[0022] 基于所述训练完成的循环神经网络,对所述第二数据进行处理,得到至少一个第三序列;其中,所述至少一个第三序列,包括所述数据库的所述第二时刻和/或所述第二时段的所述至少一个运行参数值;

[0023] 基于所述至少一个第三序列和所述至少一个第二序列,得到所述第一信息。

[0024] 在一些实施方式中,所述基于所述至少一个第三序列和所述至少一个第二序列,得到所述第一信息,包括:

[0025] 对所述至少一个第二序列进行聚类,得到至少一个类簇;

[0026] 基于所述至少一个类簇,对所述至少一个第三序列进行处理,得到所述第一信息。

[0027] 在一些实施方式中,所述基于所述训练完成的循环神经网络,对所述第二数据进行处理,得到至少一个第三序列,包括:

[0028] 将所述第二数据输入至所述训练完成的循环神经网络,得到第三数据;

[0029] 基于所述信息分类规则,对所述第三数据进行处理,得到至少一个第三序列。

[0030] 本申请还提供了一种信息处理系统,所述系统包括:处理器、存储器和通信总线;其中,所述通信总线用于实现处理器和存储器之间的通信连接;

[0031] 所述处理器用于执行存储器中的信息处理方法的程序,以实现以下步骤:

[0032] 获取第一数据;其中,所述第一数据,包括数据库第一时刻的至少一种运行状态信息和第一时段的所述至少一个运行参数信息;其中,所述第一时刻,用于表示所述数据库历史运行过程中的时刻;所述第一时段,用于表示所述第一时刻之前预设长度的时间段;

[0033] 获取第二数据;其中,所述第二数据,用于表示所述数据库当前时刻的所述至少一个运行参数信息;

[0034] 基于所述第一数据和所述第二数据,得到第一信息;其中,所述第一信息,用于表示所述数据库第二时刻和/或第二时段的运行状态信息;其中,所述第二时刻,用于表示所述数据库运行的未来时刻;所述第二时段,用于表示所述数据库运行的预设长度的未来时段。

[0035] 本申请还提供了一种信息处理设备,所述信息处理设备包括:获取模块和处理模块;其中:

[0036] 所述获取模块,用于获取第一数据和第二数据;其中,所述第一数据,包括数据库第一时刻的至少一种运行状态信息和第一时段的所述至少一个运行参数;其中,所述第一时刻,用于表示所述数据库历史运行过程中的时刻;所述第一时段,用于表示所述第一时刻之前预设长度的时间段;所述第二数据,用于表示所述数据库当前时刻的所述至少一个运行参数;

[0037] 所述处理模块,用于基于所述第一数据和所述第二数据,得到第一信息;其中,所述第一信息,用于表示所述数据库第二时刻和/或第二时段的运行状态信息;其中,所述第二时刻,用于表示所述数据库运行的未来时刻;所述第二时段,用于表示所述数据库运行的预设长度的未来时段。

[0038] 本申请还提供了一种计算机可读存储介质,所述计算机可读存储介质存储有一个或者多个程序,所述一个或者多个程序可被一个或者多个处理器执行,以实现以上任一项所述的信息处理方法的步骤。

[0039] 本申请提供的信息处理方法,首先获取表征数据库历史运行过程中的第一时刻的至少一种运行状态信息,还要获取表示第一时刻之前预设长度时间段的至少一个运行参数,然后获取表示数据库当前时刻的至少一个运行参数的第二数据,再基于第一数据和第二数据,得到表示数据库的未来时刻的运行状态信息。由此,本申请所提供的信息处理方法,可以根据数据库历史运行过程当中的运行状态信息和至少一个运行参数,结合当前时刻的至少一个运行参数的信息,预测数据库未来时刻的运行状态信息,从而解决了相关技术在数据库尚未出现异常时,无法预测数据库在未来的运行中是否会出现异常的问题。

附图说明

[0040] 图1为本申请提供的第一种信息处理方法的流程示意图;

[0041] 图2为本申请提供的第二种信息处理方法的流程示意图;

[0042] 图3为本申请提供的信息处理方法的具体实现流程图;

[0043] 图4为本申请提供的一种信息处理系统的流程示意图;

[0044] 图5为本申请提供的一种信息处理设备的结构图。

具体实施方式

[0045] 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。

[0046] 应当理解,此处所描述的具体实施例仅仅用以解释本申请,并不用于限定本申请。

[0047] 本申请提供的信息处理方法,涉及电子信息技术领域,尤其设计一种信息处理方法、系统、设备以及计算机可读存储介质。

[0048] 随着信息技术的发展,互联网或云的很多数据都存储在数据库中,随着数据库存储数据量的增大、访问数据库请求数量的增多以及数据库运行时间的增长,在数据库运行过程中不可避免的会出现一些问题,比如连接数陡增、慢请求积压、响应缓慢等,这些问题都会给用户和上层业务产生不良影响。当数据库运行过程中出现以上各种问题时,需要及

时的向数据库管理员或者技术维护人员发送异常信息。

[0049] 在相关技术中,为了实时的监控数据库的运行状态,通常会借助于第三方组件来实现,比如利用telegraf组件采集数据库运行过程中的各种参数数据,再利用influxDB组件存储这些参数数据,然后利用zabbix组件进行直观的展示,比如在坐标系中用曲线的方式进行展示,极大的方便了运行维护人员的查看和分析以上各种参数数据。

[0050] 为了根据获取到的上述参数数据确定数据库是否发生异常,相关技术中还设置了与获取的上述各种参数对应的阈值,当采集到的参数数值大于对应设置的阈值时,则认为数据库当前的运行已经发生故障,此时,可以向管理员或运行维护人员发送邮件,或者是发送信息以提醒邮件的接收方、信息的接收方及时处理故障。

[0051] 为了摆脱对第三方组件的依赖,以及固定阈值带来的故障检测不灵活的缺点,相关技术中还引入了人工智能技术(Artificial Intelligence, AI),该技术通过对数据库历史运行中发生异常时的若干关键参数及其数值进行训练,得到训练完成的AI模型。然后使用训练完成的AI模型对数据库运行过程中的上述若干关键参数对应数值的进行采集,并对采集结果进行判断,从而得到数据库当前运行状态的信息报告。

[0052] 然而,相关技术中的以上两种技术方案,都是通过监控数据库当前运行过程中的各种参数,以实现数据库当前是否发生异常的判断,也就是说,相关技术中的上述方案,都是一种被动式的监控,这些方案的实现,都需要运维人员随时关注可视化页面,并对指标进行判断,对运维人员的经验依赖较为明显,且很容易出现故障漏查的情况。并且,采用AI的方式降低了误报或漏报的概率,但该方法不能断定未来故障发生的可能性。因此,相关技术中的上述方案,对于数据库运行过程中未来时刻的状态,是否为异常状态,无法做出预测和判断,从而无法实现对数据库近期可能出现的异常的预防和处理。

[0053] 基于以上问题,本申请实施例提供了一种信息处理方法,该信息处理方法,可以通过信息处理设备的处理器来实现。如图1所示,该信息处理方法包括以下步骤:

[0054] 步骤101、获取第一数据。

[0055] 其中,第一数据,包括数据库第一时刻的至少一种运行状态信息和第一时段的至少一个运行参数;其中,第一时刻,用于表示数据库历史运行过程中的时刻;第一时段,用于表示第一时刻之前预设长度的时间段。

[0056] 在一种实施方式中,数据库,可以是存储多用户的、多种类型数据的数据库。

[0057] 在一种实施方式中,数据库,可以是存储公共数据的数据库,比如一些新闻网站对应的数据。

[0058] 在一种实施方式中,数据库,可以是存储私有数据的数据库,比如存储某一公司或组织自身数据,但不对外开放的数据库。

[0059] 在一种实施方式中,数据库,可以是数据存储时广泛采用的数据库,比如Oracle数据库、Mysql数据库等。

[0060] 在一种实施方式中,第一时刻,可以用于表示数据库历史运行过程中的某一个历史时刻。

[0061] 在一种实施方式中,第一时刻,可以用于表示数据库历史运行过程中的多个历史时刻。

[0062] 在一种实施方式中,第一时段,可以用于表示第一时刻之前固定长度的时间段,比

如5分钟。

[0063] 在一种实施方式中,第一时段,可以用于表示第一时刻之前非固定长度的时间段,比如第K个第一时段,对应的是第K个第一时刻之前长度为K1的时间段,其中,K为大于1的整数,K1为大于1的数字,单位可以为分钟。

[0064] 在一种实施方式中,第一时段,可以是根据数据库分析的需要而设定长度的第一时刻之前的时间段。

[0065] 在一种实施方式中,第一时段,可以是根据数据库存储数据类型的不同,而设定的第一时刻之前对应长度的时间段。比如,数据库中存储有第K类型的数据,则设定所有与第一类型数据操作相关的第一时段为:第一时刻之前长度为K2的时间段;其中,K2为大于1的数字,单位可以为分钟。

[0066] 在一种实施方式中,第一时段,可以是根据数据库种类的不同,而设定的第一时刻之前对应长度的时间段,比如,对于第K类型的数据库,则为其设定的第一时段为:第一时刻之前长度为K3的时间段;其中,K3为大于1的数字,单位为分钟。

[0067] 在一种实施方式中,第一时段,可以是根据数据库数据访问负载的不同,而设定的第一时刻之前对应长度的时间段,比如,对于第K个数据库,针对访问负载不同的变化区间,设定不同对应长度的时间段。

[0068] 在一种实施方式中,第一时段,可以是根据数据库运行时段的不同,而设定的第一时刻之前对应长度的时间段,比如对于第K个数据库,针对一天、一周、一个月或一个季度、一年当中不同的时段,而设定不同对应长度的时间段。

[0069] 在一种实施方式中,第一时刻的至少一种运行状态信息,可以表示第一时刻时,数据库运行状态为正常或运行状态为异常。

[0070] 在一种实施方式中,第一时刻的至少一种运行状态信息,可以表示第一时刻时,数据库运行状态为数据访问的负载饱和与否的状态。

[0071] 在一种实施方式中,第一时刻的至少一种运行状态信息,可以表示第一时刻时,数据库对任一类数据访问的相应速度是否卡顿的状态。

[0072] 在一种实施方式中,第一时刻的至少一种运行状态信息,可以表示第一还时刻时,数据库对指定类别的数据访问的速度是否卡顿的状态。

[0073] 在一种实施方式中,第一时刻的至少一种运行状态信息,还可以包括第一时刻时至少一种运行状态信息对应的数据库的标识信息。

[0074] 在一种实施方式中,第一时刻的至少一种运行状态信息,除了包括第一时刻数据库运行的具体状态的信息和数据库的标识信息之外,还可以包括与各种具体状态对应的可以表征数据库运行状态的至少一个参数信息。

[0075] 在一种实施方式中,第一时段的至少一个运行参数,可以是用于表征数据库运行状态的若干关键运行参数。

[0076] 在一种实施方式中,第一时段的至少一个运行参数,可以是根据数据库运行状态的不同而对应设定的若干关键运行参数。

[0077] 在一种实施方式中,第一时段的至少一个运行参数,可以是根据数据库运行时段的不同而对应设定的若干关键运行参数。

[0078] 在一种实施方式中,第一时段的至少一个运行参数,可以是根据数据库中存储数

据的不同,而对应设定的若干关键运行参数。

[0079] 在一种实施方式中,第一时段的至少一个运行参数,可以是根据数据库类型的不同,而对应设定的若干关键运行参数。

[0080] 在一种实施方式中,为了能够充分而全面的反映出数据运行状态,第一时段的至少一个运行参数,可以包括至少三个关键的运行参数。

[0081] 在一种实施方式中,第一时刻的至少一种运行状态信息,与第一时段的至少一个运行参数,是一一对应的关系,比如第M种运行状态信息,与第M组至少一个运行参数相对应。其中,M为大于或等于1的整数。

[0082] 步骤102、获取第二数据。

[0083] 其中,第二数据,用于表示数据库当前时刻的至少一个运行参数。

[0084] 在一种实施方式中,第二数据,可以用于表示数据库当前时刻运行过程中,能够反映数据库运行状态的若干关键运行参数。

[0085] 在一种实施方式中,第二数据,可以是当前时刻,表示数据库对某一指定数据访问对应的运行状态的若干关键的运行参数。

[0086] 在一种实施方式中,第二数据,可以是当前时刻,与数据库的类型对应的、能够反映数据库运行状态的运行参数。

[0087] 在一种实施方式中,第二数据,可以是当前时刻,与数据库的运行时段对应的、能够反映数据库运行状态的参数。

[0088] 在一种实施方式中,第二数据,可以是能够反映数据库当前运行状态的、并且与第一时段的至少一个运行参数对应的参数。

[0089] 需要说明的是,为了能够高效的对获取到的第二数据进行处理,在本申请实施例提供的信息处理方法的实现过程中,第一数据和第二数据获取操作,是依次获取的关系。

[0090] 步骤103、基于第一数据和第二数据,得到第一信息。

[0091] 其中,第一信息,用于表示数据库的第二时刻和/或第二时段的运行状态信息;第二时刻,用于表示数据库运行的未来时刻;第二时段,用于表示数据库运行的预设长度的未来时段。

[0092] 在一种实施方式中,第一信息,可以用于表示当前时刻相邻的下一时刻和/或下一时段内,数据库的运行状态信息。

[0093] 在一种实施方式中,第一信息,可以用于表示与当前时刻间隔的一段时间之后的第二时刻和/或第二时段内,数据库的运行状态信息,比如,当前时刻的5分钟之后,数据库的运行状态信息;和/或,当前时刻的五分钟之后,三分钟的时间长度内数据库的运行状态信息。

[0094] 在一种实施方式中,第一信息,可以是根据数据库的不同,而设定的与当前时刻对应时段之后,的第二时刻和/或第二时段内,数据库的运行状态信息,比如,第K种数据库的当前时刻的K4时段之后,和/或,K4时段之后,长度为M1的时段内的数据库的运行状态信息;其中,K4为大于0的数字,单位为分钟;M1为大于0的数字,单位可以为分钟。

[0095] 在一种实施方式中,第一信息,可以是在当前时刻的基础上,根据数据库运行时段的的不同,而设定的与运行时段对应的延迟时段后的第二时刻和/或第二时段内的运行状态信息。

[0096] 在一种实施方式中,第一信息,可以包括数据库的标识信息。

[0097] 在一种实施方式中,第一信息,可以是与第一数据对应的、用于表示数据库的第二时刻时、和/或第二时段内的运行状态信息。

[0098] 在实际应用中,上述处理器可以为专用集成电路(Application Specific Integrated Circuit,ASIC)、数字信号处理器(digital signal processor),DSP)、可编程逻辑器件(Programmable Logic Device,PLD)、片上可编程逻辑门阵列(Field Programmable Gate Array,FPGA)、中央处理器(central processing unit,CPU)、控制器、微控制器、微处理器中的至少一种。

[0099] 本申请实施例提供的信息处理方法,首先获取表征数据库历史运行过程中的第一时刻的至少一种运行状态信息,还要获取表示第一时刻之前预设长度时间段的至少一个运行参数,然后获取表示数据库当前时刻的至少一个运行参数的第二数据,再基于第一数据和第二数据,得到表示数据库的未来时刻的运行状态信息。由此,本申请实施例所提供的信息处理方法,可以根据数据库历史运行过程当中的运行状态信息和至少一个运行参数,结合当前时刻的至少一个运行参数的信息,预测数据库未来时刻和/或未来时段的运行状态信息,从而解决了相关技术中在数据库尚未出现异常时,无法预测数据库在未来的运行中是否会出现异常的问题。

[0100] 基于前述实施例,本申请实施例提供了一种信息处理方法,如图2所示,该信息处理方法包括以下步骤:

[0101] 步骤201、获取第一数据。

[0102] 其中,第一数据,包括数据库第一时刻的至少一种运行状态信息和第一时段的至少一个运行参数;其中,第一时刻,用于表示数据库历史运行过程中的时刻;第一时段,用于表示第一时刻之前预设长度的时间段。

[0103] 步骤202、获取第二数据。

[0104] 其中,第二数据,用于表示数据库当前时刻的至少一个运行参数。

[0105] 步骤203、获取信息分类规则。

[0106] 在步骤203中,信息分类规则,可以是对第二数据进行分类的规则。

[0107] 在一种实施方式中,信息分类规则,可以是根据第二数据中所包含的运行参数的类型相对应的分类规则。

[0108] 在一种实施方式中,信息分类规则,可以是根据第一数据中所包含的运行参数的类型相对应的分类规则。

[0109] 在一种实施方式中,信息分类规则,可以是用于从第二数据或第二数据所包含的多个运行参数中、筛选出指定几种类型的参数的分类规则。

[0110] 在一种实施方式中,信息分类规则,可以根据数据库类型的不同而不同。

[0111] 在一种实施方式中,信息分类规则,可以根据数据库运行时段的的不同而不同。

[0112] 在一种实施方式中,信息分类规则,可以是根据数据库对不同类型的数据访问操作而对应的不同的信息分类规则。

[0113] 步骤204、基于信息分类规则,对第一数据中的至少一个运行参数进行处理,得到至少一个第一序列。

[0114] 其中,第一序列,用于表示在第一时段至少一个运行参数的状态变化的序列。

[0115] 在一种实施方式中,第一序列,可以用于表示数据库在第一时段内的、第一数据中至少一个运行参数的变化量的序列。

[0116] 在一种实施方式中,第一序列,可以用于表示数据库在第一时段内的、第一数据中至少一个运行参数变化量的百分比的序列。

[0117] 在一种实施方式中,第一序列,可以是首先获取数据库在第一时段内的、第一数据中至少一个运行参数的变化量,再将该变化量与预设的第一数据中至少一个运行参数的变化区间进行匹配,得到的一系列的变化量匹配结果。

[0118] 在一种实施方式中,第一序列,可以是第一数据中指定的若干运行参数的状态变化的序列。

[0119] 在一种实施方式中,基于信息分类规则,对第一数据中的至少一个运行参数进行处理,得到至少一个第一序列,可以是基于信息分类规则,首先对第一数据中的至少一个运行参数进行分类,再对分类结果进行变化量统计,然后基于变化量统计的结果与分类结果,得到至少一个第一序列。

[0120] 在一种实施方式中,基于信息分类规则,对第一数据中的至少一个运行参数进行处理,得到至少一个第一序列,可以是基于信息分类规则,将第一数据中的至少一个运行参数进行分类,对分类后的每个运行参数的变化量进行编码,然后基于分类结果和编码结果,得到至少一个第一序列。

[0121] 示例性的,步骤204,可以通过步骤A1-A2来实现:

[0122] 步骤A1、基于信息分类规则,得到第一分类规则和第二分类规则。

[0123] 其中,第一分类规则,用于表示第一数据中的至少一个运行参数的类别划分规则;第二分类规则,用于表示第一数据中的至少一个运行参数的参数值的区间划分规则。

[0124] 在一种实施方式中,第一分类规则,可以是对第一数据的至少一个运行参数中、指定的几个运行参数进行类别划分的规则。

[0125] 在一种实施方式中,第一分类规则,可以是将第一数据的每一运行参数划分至不同的类别的规则。

[0126] 在一种实施方式中,第一分类规则,可以是根据第一数据中每一运行参数的参数值将每一运行参数划分为不同类别的规则。

[0127] 在一种实施方式中,第二分类规则,可以是对第一数据的至少一个运行参数中、指定的若干运行参数对应的参数值进行区间划分的规则。

[0128] 在一种实施方式中,第二分类规则,可以是将第一数据的至少一个运行参数的参数值,与预设的参数区间之间进行匹配的规则。

[0129] 在一种实施方式中,第二分类规则,可以是将第一数据的至少一个运行参数中任一运行参数的参数值进行组合运算后,根据组合运算的结果进行区间划分的规则。

[0130] 步骤A2、基于第一分类规则和第二分类规则,对第一数据中的运行参数进行处理,得到至少一个第一序列。

[0131] 在一种实施方式中,基于第一分类规则和第二分类规则,对第一数据中的运行参数进行处理,得到至少一个第一序列,可以是基于第一分类规则,对第一数据中的至少一个运行参数进行类别划分,基于第二分类规则,将运行参数类别划分后的得到的类别划分结果中每一运行参数的参数值,进行区间划分,并类别划分结果与对应的区间划分结果进行

组合,得到至少一个第一序列。

[0132] 在一种实施方式中,基于第一分类规则和第二分类规则,对第一数据中的运行参数进行处理,得到至少一个第一序列,可以是对第一数据的运行参数中指定运行参数进行类别划分,基于第二分类规则,将类别划分的结果中对应的运行参数的参数值进行区间划分,然后将类别划分的结果与区间划分的规则进行组合,得到至少一个第一序列。

[0133] 在一种实施方式中,当第一数据中的运行参数属于同一类型时,则不需要第一分类规则对运行参数进行划分,可以直接基于第二分类参数进行区间划分,从而得到至少一个第一序列。

[0134] 步骤205、对至少一种运行状态信息和至少一个第一序列进行组合,得到至少一个第二序列。

[0135] 在一种实施方式中,至少一种运行状态信息,是从第一数据中获取到的。

[0136] 在一种实施方式中,对至少一种运行状态信息和至少一个第一序列进行组合,得到至少一个第二序列,可以是至少一种运行状态信息与得到的每一第一序列对应组合,从而得到至少一个第二序列。

[0137] 在一种实施方式中,对至少一种运行状态信息和至少一个第一序列进行组合,得到至少一个第二序列,可以是至少一种运行状态信息中选定若干运行状态信息,并相应的从第一序列中选择对应的序列,然后将选定的若干运行状态信息和对应的序列进行组合,从而得到至少一个第二序列。

[0138] 步骤206、基于至少一个第二序列,对第二数据进行处理,得到第一信息。

[0139] 在一种实施方式中,至少一个第二序列中,包括了第一序列中第一时段内至少一个运行参数的状态变化的信息。

[0140] 在一种实施方式中,基于至少一个第二序列,对第二数据进行处理,得到第一信息,可以是基于至少一个第二序列中包括的第一时段内至少一个运行参数的状态变化的信息,而确定第二数据的变化轨迹,并从该变化轨迹中,得到第一信息。

[0141] 在一种实施方式中,基于至少一个第二序列,对第二数据进行处理,得到第一信息,可以是至少一个第二序列进行参数和参数值的双重匹配,从而得到第二数据对应的变化轨迹,再基于该变化轨迹,得到第一信息。

[0142] 示例性的,步骤206,可以通过步骤B1-步骤B2来实现:

[0143] 步骤B1、获取训练完成的循环神经网络。

[0144] 其中,循环神经网络,用于预测数据库的第二时刻或第二时段的至少一个运行状态参数值。

[0145] 在一种实施方式中,循环神经网络,可以是用于预测数据库的未来时刻的、至少一个运行状态参数中若干指定运行参数对应的参数值的神经网络。

[0146] 在一种实施方式中,循环神经网络,可以是时间循环神经网络。

[0147] 在一种实施方式中,循环神经网络,可以是长短期记忆神经网络(Long Short-Term Memory,LSTM)。

[0148] 在一种实施方式中,循环神经网络,可以是门控循环单元神经网络(Gated Recurrent Unit,GRU)。

[0149] 在一种实施方式中,训练完成的循环神经网络,可以是基于训练样本数据,通过以

下至少之一的方式进行训练得到的：无监督的学习方式、有监督的学习方式、半监督的学习方式。

[0150] 示例性的，步骤B1，可以通过步骤B101来实现：

[0151] 步骤B101、基于第一数据，对循环神经网络进行训练，得到训练完成的循环神经网络。

[0152] 在一种实施方式中，基于第一数据，对循环神经网络进行训练，得到训练完成的循环神经网络，可以是将第一数据中第一时段的至少一个运行参数对应的参数值输入至循环神经网络，并将循环神经网络的输出与第一数据中的至少一种运行状态信息进行匹配，然后根据匹配的结果，调整循环神经网络的各种参数。

[0153] 在一种实施方案中，基于第一数据，对循环神经网络进行训练，得到训练完成的循环神经网络，可以是将第一数据中第一时段内的至少一个运行参数输入至循环神经网络，从而使得该循环神经网络能够根据至少一个运行参数的参数值，调整循环神经网络的每一参数，从而使得循环神经网络能够获取到第一时段内至少一个运行参数的变化趋势。

[0154] 步骤B2、基于训练完成的循环神经网络和至少一个第二序列，对第二数据进行处理，得到第一信息。

[0155] 在一种实施方式中，基于训练完成的循环神经网络和至少一个第二序列，对第二数据进行处理，得到第一信息，可以是将第二数据输入至训练完成的循环神经网络，得到处理结果，其中，处理结果，用于表示与数据库未来时刻的运行状态相关联的至少一个状态参数；然后基于至少一个第二序列对以上处理结果进行分析，得到第一信息。

[0156] 在一种实施方式中，基于训练完成的循环神经网络和至少一个第二序列，对第二数据进行处理，得到第一信息，可以是将至少一个第二序列和第二数据同时输入至循环神经网络中，根据循环神经网络的输出结果，得到第一信息。

[0157] 示例性的，步骤B2还可以通过步骤C1-步骤C2来实现：

[0158] 步骤C1、基于训练完成的循环神经网络，对第二数据进行处理，得到至少一个第三序列。

[0159] 其中，至少一个第三序列，包括数据库第二时刻和/或第二时段的至少一个运行参数值。

[0160] 在一种实施方式中，至少一个第三序列，可以包括数据库未来时刻的、至少一个运行参数信息中的、若干指定的运行参数对应的运行参数值。

[0161] 相应地，基于训练完成的循环神经网络，对第二数据进行处理，得到至少一个第三序列，可以是将第二数据输入至训练完成的循环神经网络中，得到包括数据库未来时刻的、至少一个运行参数中的若干指定的运行参数对应的运行参数值。

[0162] 在一种实施方式中，至少一个第三序列，可以用于表示数据库未来的某一指定时刻的至少一个运行参数对应的运行参数值。

[0163] 相应地，基于训练完成的循环神经网络，对第二数据进行处理，得到至少一个第三序列，可以是将第二数据输入至训练完成的循环神经网络中，得到用于表示数据库未来的某一指定时刻的、至少一个运行参数对应的运行参数值。

[0164] 在一种实施方式中，至少一个第三序列，可以用于表示数据库未来的任一时段内的至少一个运行参数对应的运行参数值。

[0165] 相应地,基于训练完成的循环神经网络,对第二数据进行处理,得到至少一个第三序列,可以是第二数据输入至训练完成的循环神经网络中,得到用于表示数据库未来的任一时段内的、至少一个运行参数对应的运行参数值。

[0166] 在一种实施方式中,至少一个第三序列,可以用于表示数据库未来的某一指定时段内的至少一个运行参数对应的运行参数值。

[0167] 相应地,基于训练完成的循环神经网络,对第二数据进行处理,得到至少一个第三序列,可以是第二数据输入至训练完成的循环神经网络中,得到用于表示数据库未来的某一指定时段内的、至少一个运行参数对应的运行参数值。

[0168] 在一种实施方式中,至少一个第三序列,可以用于表示数据库未来的某一指定时段内的、至少一个运行参数中的若干指定运行参数的值。

[0169] 相应地,基于训练完成的循环神经网络,对第二数据进行处理,得到至少一个第三序列,可以是第二数据输入至训练完成的循环神经网络中,得到用于表示数据库未来的某一指定时段内的、至少一个运行参数信息中的若干指定的运行参数对应的运行参数值。

[0170] 示例性的,步骤C1,还可以通过步骤D1-步骤D2来实现:

[0171] 步骤D1、将第二数据输入至训练完成的循环神经网络,得到第三数据。

[0172] 在一种实施方式中,可以是第二数据输入至训练完成的循环神经网络,使得训练完成的循环神经网络可以基于第二数据进行预测,得到用于表示数据库的未来时刻运行状态的至少一个运行参数值。

[0173] 在一种实施方式中,第三数据,还可以包括数据库标识信息、未来时刻的时间信息以及第三数据的发展趋势信息。

[0174] 步骤D2、基于信息分类规则,对第三数据进行处理,得到至少一个第三序列。

[0175] 在一种实施方式中,至少一个第三序列,可以是根据信息分类规则中的第一分类规则和/或第二分类规则,对第三数据进行处理得到的。

[0176] 具体地,当第三数据中仅包括一种类型的运行参数时,可以直接基于第二分类规则,对该运行参数对应的参数值进行区间划分。

[0177] 当第三数据中包括至少两种类型的运行参数时,可以首先基于第一分类规则,对运行参数进行分类,然后基于第二分类规则,对各种类型的运行参数对应的参数值进行区间划分。

[0178] 步骤C2、基于至少一个第三序列和至少一个第二序列,得到第一信息。

[0179] 在一种实施方式中,第一信息,可以是至少一个第三序列与至少一个第二序列进行匹配得到的。

[0180] 示例性的,步骤C2,可以通过步骤E1-步骤E2来实现。

[0181] 步骤E1、对至少一个第二序列进行聚类,得到至少一个类簇。

[0182] 在一种实施方式中,类簇,用于表示至少一个第二序列中,具备相同特征的第二序列的集合。

[0183] 在一种实施方式中,从至少一个第二序列进行聚类,得到的至少一个类簇中,不同类簇中的第二序列的特征彼此相异。

[0184] 在一种实施方式中,对至少一个第二序列进行聚类,可以通过基于密度对噪声鲁棒的空间聚类算法(Density-Based Spatial Clustering of Applications with

Noise,DBSCAN)来实现的。

[0185] 在一种实施方式中,对至少一个第二序列进行聚类,得到至少一个类簇,可以从第一数量的第二序列,得到第二数量的类簇,其中,第二数量要小于第一数量,在一些实施方式中,第二数量可以远远小于第一数量,由此,可以实现数据量的显著减少。

[0186] 步骤E2、基于至少一个类簇,对至少一个第三序列进行处理,得到第一信息。

[0187] 在一种实施方式中,基于至少一个类簇,对至少一个第三序列进行处理,得到第一信息,可以是至少一个类簇中的每一类簇,与至少一个第三序列中的每一序列进行匹配,从而得到第一信息。

[0188] 在一种实施方式中,基于至少一个类簇,对至少一个第三序列进行处理,得到第一信息,可以是首先对第三序列进行分析,再将分析结果与至少一个类簇进行特征匹配,从而得到第一信息。

[0189] 在一种实施方式中,基于至少一个类簇,对至少一个第三序列进行处理,得到第一信息,可以将至少一个类簇中的任一类簇,与至少一个第三序列中的任一第三序列进行匹配,得到第P个类簇与第L个第三序列匹配成功的结果,然后根据匹配成功的结果得到第一信息;其中P和L,均为大于0的整数。

[0190] 在一种实施方式中,第一信息,还可以是至少一个类簇中的任一类簇,与至少一个第三序列中的任一序列匹配失败的结果。此时,需要基于当前的第三序列,和/或,与第三序列对应的第二数据,对第一数据进行更新,并根据更新的结果更新至少一个第二序列,由此,可以得到数据覆盖面更广、特征更加完善的新的类簇。之后,可以根据新的类簇,对至少一个第三序列进行处理,得到第一信息。

[0191] 具体地,图3为本申请实施例所提供的信息处理方法的具体实现流程图。在图3中所示的信息处理方法的具体实现流程图中,主要包括状态关联计算模块、时序指标预测模块、状态关联知识库、状态实时预测模块,还包括第二数据获取部分以及故障清单部分。

[0192] 在图3中,实例1、实例2以及实例3对应的是不同的数据库的实例,其中,实例1-3可以是同一类型的数据库的实例,也可以是彼此类型互不相同的数据库的实例。

[0193] 在图3中,第二数据,包括从实例1-3中实时获取到的当前时刻的运行状态信息以及至少一个运行参数,即图3中对应的当前参数。

[0194] 在图3中,第一数据,包括从实例1-3中获取到的第一时刻的至少一个运行状态信息和第一时段的至少一个运行参数。

[0195] 在图3中,状态关联计算模块,用于获取第一数据,示例性的,还用于根据需要从第一数据中提取和筛选出需要分析的目标参数,比如与数据库异常运行状态相关联的若干关键参数信息,以及与该若干关键参数信息对应的异常信息。具体地,比如,实例1在第一时刻发生的数据请求无响应问题,或,实例1在第一时刻之前的第一时段内的与该数据请求无响应问题相关的至少一个运行参数及其参数值。其中,第一时刻,比如可以是2019年10月1日23点35分20秒,第一时段,可以是2019年10月1日23点30分20秒-2019年10月1日23点35分20秒之间的时间段。示例性的,第一时段,可以是第一时刻之前的任意时间长度的时间段,比如三分钟、四分钟等。示例性的,至少一个运行参数,比如可以是实时连接数(Real connection Count,RC)、慢连接数(Slow quest Count,SC)以及响应时间(Response Time,RT)。数据库的运行状态,可以是数据库运行过程中的异常状态。

[0196] 在图3中,状态关联计算模块,还用于根据获取的第一数据得到至少一个第一序列。具体地,在至少一个运行参数包括RC、SC和RT的情况下,至少一个第一序列中可以包括RC、SC以及RT的参数值。并且,第一序列中所包含的以上各种参数对应参数值所覆盖的时间长度,可以与第一时段匹配。对每一序列内部的至少一个运行参数值而言,可以是数据采集时的瞬时值,也可以是数据采集时预设时间长度范围内的平均值,比如在20秒内的RC的平均值。

[0197] 示例性地,第一序列,可以是以下格式:RC/SC/RT,将第一数据中的运行状态信息与第一序列进行组合,则可以得到第二序列,第二序列的格式如下:状态名称/状态发生位置/状态发生时间/RC/SC/RT,上述第二序列中状态名称,可以用于表示某一具体的状态对应的名称,比如RC过少;状态发生位置,可以表示状态名称对应的状态发生的数据库实例,比如实例1;状态发生时间,可以表示状态名称对应的状态实际发生的时间,比如,XXX年X月XX日X时X分,RC/SC/RT表示状态发生时间对应的能够表征RC、SC、RT的参数状态。在第一序列中包含的状态为故障状态时,第二序列可以包括以下内容:故障名称/故障位置/故障发生时间/RC/SC/RT。其中,故障名称,可以是故障的具体名称,比如无响应;状态发生位置,可以是故障发生的数据库实例,比如Mysql数据库实例2,故障发生时间,可以是精确至秒或者毫秒级别的时间,比如2019年12月25日23点34份15秒,RC/SC/RT可以用于表示故障发生时表征RC/SC/RT的参数的状态。

[0198] 在图3中,状态关联计算模块,还用于基于信息分类规则,对得到的至少一个第二序列进行分类,并对每一序列中的参数区间进行区间划分,得到至少一个信息扩展后的第二序列。比如,从第二序列中获取状态名称,比如响应过长,获取以上状态发生的数据库标识,比如Mysql数据库的实例2,还可以获取状态出现的时间,比如2019年12月25日23点34份15秒。具体地,在至少一个运行参数为RC、SC以及RT的情况下,可以对以上三个参数的区间进行划分,从而将各个参数对应的参数值的变化趋势反应出来。

[0199] 示例性的,状态关联计算模块,还可以将第二序列中的每个参数扩充为两个对应的子参数Y和Z,其中,子参数Y用于表示每个参数值的区间值,子参数Z为每个参数对应的参数值的变化率的区间值。并且,为了更好的对各个参数的参数值进行区间划分,分别对子参数Y对应值和子参数Z对应值进行区间划分,示例性地,对于子参数Y的区间划分,可以将子参数Y参数值的最大值与最小值之间的区间等分为五个区间,例如,子参数Y,可以划分为以下5个区间:EL很低/SL偏低/C普通/SH偏高/EH很高;示例性地,对于子参数Z的区间划分,可以将第m时刻子参数Z的值与第m-1时刻的子参数Z的至进行差分计算,以得到第m时刻子参数Z与第m-1时刻子参数Z的变化比例,其中,m为大于1的整数,比如,子参数Z,可以划分为以下5个对应的区间:按照不足20-20%/230-(20%-30%)/370(30%-70%)/710-大于70%,并且,考虑到还需要体现出变化趋势,示例性地,还可以增加一个维度的子参数X,子参数X的取值可以为A上升或R下降,由此,可以得到A20/A230/A370/A710/R20/R230/R370/R710八个取值区间,其中,A表示增加,R表示降低,上述八个取值区间分别表示增加20%/增加20%-30%/增加30%-70%/增加大于70%/降低20%/降低20-30%/降低30%-70%/降低大于70%。

[0200] 状态关联计算模块,还可以在以上8个取值区间前面加上参数编号,则可以得到比如:RC-EH,代表实时连接数指标当前很高;RT-R230,代表响应时间指标降低处于20%-30%

之间。

[0201] 经过状态关联计算模块的以上处理,通过第一数据,可以得到格式如下的至少一个信息扩展后的第二序列:故障名称/故障位置/故障时间/实时连接数当前指标值区间/实时连接数变化率区间/慢连接数量/慢连接数量变化率区间/响应时间当前指标值区间/响应时间变化率区间,共9个字段,示例如下:响应时间过长/南方CRM5号数据库3号实例/20191010163429/RC-EH/RC-A20/SC-EL/SC-R710/RT-C/RT-R370,以上示例中各个字段的含义如下:南方CRM5号数据库3号实例在2019年10月10日16时34分29秒出现了响应过长时间现象,并且,对应的在第一时段内出现了RC过高,并增长了20%,SC过低,并下降了大于70%,RT处于一般水平,并下降了30%-70%。

[0202] 另外,状态关联计算模块,还用于对至少一个信息扩展后的第二序列进行聚类,得到至少一个类簇。经过状态关联计算模块处理得到的至少一个第二序列的数量,可能是非常大的,这样大的数据量,对于后续的数据处理,也显得便捷性不足,并且,对第二序列进行聚类得到的类簇中包括了至少一个第二序列中足够的运行参数特征信息,还可以大大降低数据量。

[0203] 经过状态关联计算模块处理后得到的多个类簇中的每一类簇的特征明显,并且可以具有广泛的代表意义,比如,类簇一:包含123条故障,其特征可大致表述为:数据库响应慢-RC较高-RC增加30%到70%-RT较高-RT升高70%以上;类簇二:包含234条故障,其特征可大致表述为:RC过高-SC较低-SC数降低20%以下-RT很高-RT升高70%以上。并且,在类簇中所包含的第二序列的数量越多,则该类簇的有效性越明显,在第二序列表示故障序列的条件下,在各个类簇中包含第二序列的数量越多的类簇,可以是对后续数据处理越有效的类簇。

[0204] 状态关联计算模块,还用于将得到的各个类簇存储于状态关联知识库,示例性地,可以设定第二序列的数量大于P的类簇为有效类簇,P可以根据实际情况设定,P可以为大于0的整数,比如20,并且可以把这些有效类簇存储在状态关联知识库中。在至少一个运行参数代表的是故障发生时的关键参数信息时,状态关联知识库可以是故障指标关联知识库(Fault Relationship Database, FRDB),相应地,可以将与故障对应的有效类簇存储在故障指标关联知识库中,存储的规则如下:类簇编号/类簇包含故障数/三个参数(每个参数(RC/SC/RT)当前参数值区间编码/参数值变化区间编码两字段)/故障名称/故障位置,一共10个字段。

[0205] 在图3中,时序指标预测模块,用于获取数据库当前时刻的至少一个运行参数,即第二数据,还用于通过训练完成的循环神经网络对当前时刻的至少一个运行参数进行处理,以预测第二时刻和/或第二时段的至少一个运行参数的参数值。示例性的,循环神经网络,可以是LSTM。

[0206] 在图3中,状态实时预测模块,用于通过上述将第一数据处理得到至少一个信息扩展后的第二序列的办法,将预测得到的第二时刻和/或第二时段的至少一个运行参数的参数值进行处理,得到至少一个第三序列。并且,至少一个第三序列中各个参数、参数值对应区间以及升降趋势等,与至少一个第二序列保持一致。

[0207] 在至少一个运行参数表示的是故障发生时的数据库的运行参数时,至少一个第三序列,可以是第二时刻和/或第二时段的故障发生时各种运行参数、参数值对应区间以及升

降趋势等的信息集合。

[0208] 状态实时预测模块,还用于从状态关联知识库中获取各种类簇,并将至少一个第三序列与各个类簇进行匹配,从而确定至少一个第三序列所属类簇。示例性的,在至少一个第三序列代表故障发生时各种参数信息的情况下,状态实时预测模块从FRDB中获取类簇,并将至少一个关联信息与各个类簇进行匹配。当匹配成功时,则输出故障清单,故障清单,可以包括至少一个第三序列中所包含的所有信息,也可以是对至少一个第三序列中的信息进行提炼,形成更直观的信息,比如,“Mysql数据库的实例3,将会在2020年1月15日16点23分38秒出现RC相应过慢的故障,请及时关注”。此外,还可以通过邮件或者信息的形式将故障清单中的数据发送至数据库管理员,或者发送至数据的运行维护人员。

[0209] 由此,本申请实施例提供的信息处理方法,获取包括数据库历史运行过程中时刻的第一时刻至少一种运行状态信息,获取数据库第一时刻之前预设长度的时间段即第一时段的至少一个运行参数的第一参数,并且,获取包括数据库当前时刻的至少一个运行参数的第二参数,然后根据信息分类规则,至少一个运行参数进行处理,得到至少一个第二序列,因而得到的至少一个第二序列相对于第一数据而言,数据信息排布更加整体、数据处理更加便捷,再基于至少一个第二序列对第二数据进行处理,得到表示数据库未来时刻和/或未来时段运行状态信息的第一信息。也就是说,在本申请实施例所提供的信息处理方法中,第一信息的获取,是基于数据库运行历史时刻中的至少一个运行参数和当前时刻的至少一个运行参数信息进行的,通过以上方式获取的第一信息,可以更贴合数据库运行状态的实际情况,即对数据库未来时刻和/或未来时段运行状态的预测,可以更加准确,从而可以提前提示数据库管理员或数据库技术维护人员提前进行可能发生的故障的干预和处理,进而可以减轻严重的或者大规模的数据库故障对各种数据业务造成的冲击。

[0210] 基于前述实施例,本申请实施例提供了一种信息处理系统3,如图4所示,该信息处理系统3包括,处理器31、存储器32和通信总线;

[0211] 其中,通信总线用于实现处理器31和存储器32之间的通信连接;

[0212] 处理器31,用于执行存储器32中的信息处理方法的程序,以实现以下步骤:

[0213] 获取第一数据;其中,第一数据,包括数据库第一时刻的至少一种运行状态信息和第一时段的至少一个运行参数信息;其中,第一时刻,用于表示数据库历史运行过程中的时刻;第一时段,用于表示第一时刻之前预设长度的时间段;

[0214] 获取第二数据;其中,第二数据,用于表示数据库当前时刻的至少一个运行参数信息;

[0215] 基于第一数据和第二数据,得到第一信息;其中,第一信息,用于表示数据库第二时刻和/或第二时段的运行状态信息;其中,第二时刻,用于表示数据库运行的未来时刻;第二时段,用于表示数据库运行的预设长度的未来时段。

[0216] 处理器31,用于执行存储器32中的信息处理方法的程序,以实现以下步骤:

[0217] 基于第一数据和第二数据,得到第一信息,包括:

[0218] 获取信息分类规则;

[0219] 基于信息分类规则,对第一数据中的至少一个运行参数进行处理,得到至少一个第一序列;其中,第一序列,用于表示在第一时段至少一个运行参数的状态变化的序列;

[0220] 对至少一种运行状态信息和至少一个第一序列进行组合,得到至少一个第二序

列;

[0221] 基于至少一个第二序列,对第二数据进行处理,得到第一信息。

[0222] 处理器31,用于执行存储器32中的信息处理方法的程序,以实现以下步骤:

[0223] 基于信息分类规则,对第一数据中的至少一个运行参数进行处理,得到至少一个第一序列,包括:

[0224] 基于信息分类规则,得到第一分类规则和第二分类规则;其中,第一分类规则,用于表示第一数据中的至少一个运行参数的类别划分规则;第二分类规则,用于表示第一数据中的至少一个运行参数的参数值的区间划分规则;

[0225] 基于第一分类规则和第二分类规则,对第一数据中的运行参数进行处理,得到至少一个第一序列。

[0226] 处理器31,用于执行存储器32中的信息处理方法的程序,以实现以下步骤:

[0227] 基于至少一个第二序列,对第二数据进行处理,得到第一信息,包括:

[0228] 获取训练完成的循环神经网络;其中,循环神经网络,用于预测数据库的第二时刻和/或第二时段的至少一个运行状态参数值;

[0229] 基于训练完成的循环神经网络和至少一个第二序列,对第二数据进行处理,得到第一信息。

[0230] 处理器31,用于执行存储器32中的信息处理方法的程序,以实现以下步骤:

[0231] 获取训练完成的循环神经网络,包括:

[0232] 基于第一数据,对循环神经网络进行训练,得到训练完成的循环神经网络。

[0233] 处理器31,用于执行存储器32中的信息处理方法的程序,以实现以下步骤:

[0234] 基于训练完成的循环神经网络和至少一个第二序列,对第二数据进行处理,得到第一信息,包括:

[0235] 基于训练完成的循环神经网络,对第二数据进行处理,得到至少一个第三序列;其中,至少一个第三序列,包括数据库第二时刻和/或第二时段的至少一个运行参数信息值;

[0236] 基于至少一个第三序列和至少一个第二序列,得到第一信息。

[0237] 处理器31,用于执行存储器32中的信息处理方法的程序,以实现以下步骤:

[0238] 基于至少一个第三序列和至少一个第二序列,得到第一信息,包括:

[0239] 对至少一个第二序列进行聚类,得到至少一个类簇;

[0240] 基于至少一个类簇,对至少一个第三序列进行处理,得到第一信息。

[0241] 处理器31,用于执行存储器32中的信息处理方法的程序,以实现以下步骤:

[0242] 基于训练完成的循环神经网络,对第二数据进行处理,得到至少一个第三序列,包括:

[0243] 将第二数据输入至训练完成的循环神经网络,得到第三数据;

[0244] 基于信息分类规则,对第三数据进行处理,得到至少一个第三序列。

[0245] 在实际应用中,上述处理器31可以为ASIC、DSP、PLD、FPGA、CPU、控制器、微控制器、微处理器中的至少一种。

[0246] 上述存储器32可以是易失性存储器(Volatile Memory),例如RAM;或者非易失性存储器(Non-volatile Memory),例如ROM,快闪存储器(Flash Memory),硬盘(Hard Disk Drive,HDD)或固态硬盘(Solid-State Drive,SSD);或者上述种类的存储器的组合,并向处

理器31提供指令和数据。

[0247] 本申请实施例提供的信息处理系统3,首先获取表征数据库历史运行过程中的第一时刻的至少一种运行状态信息,还要获取表示第一时刻之前预设长度时间段的至少一个运行参数,然后获取表示数据库当前时刻的至少一个运行参数的第二数据,再基于第一数据和第二数据,得到表示数据库的未来时刻的运行状态信息。由此,本申请实施例所提供的信息处理系统3,可以根据数据库历史运行过程当中的运行状态信息和至少一个运行参数,结合当前时刻的至少一个运行参数的信息,预测数据库未来时刻和/或未来时段的运行状态信息,从而解决了相关技术中在数据库尚未出现异常时,无法预测数据库在未来的运行中是否会出现异常的问题。

[0248] 基于前述实施例,本申请实施例提供了一种信息处理设备4,如图5所示,该信息处理设备4包括:获取模块41和处理模块42;其中,

[0249] 获取模块41,用于获取第一数据和第二数据;其中,第一数据,包括数据库第一时刻的至少一种运行状态信息和第一时段的至少一个运行参数;其中,第一时刻,用于表示数据库历史运行过程中的时刻;第一时段,用于表示第一时刻之前预设长度的时间段;第二数据,用于表示数据库当前时刻的至少一个运行参数;

[0250] 处理模块42,用于基于第一数据和第二数据,得到第一信息;其中,第一信息,用于表示数据库第二时刻和/或第二时段的运行状态信息;其中,第二时刻,用于表示数据库运行的未来时刻;第二时段,用于表示数据库运行的预设长度的未来时段。

[0251] 获取模块41,用于获取信息分类规则;

[0252] 处理模块42,用于基于信息分类规则,对第一数据中的至少一个运行参数进行处理,得到至少一个第一序列;对至少一种运行状态信息和至少一个第一序列进行组合,得到至少一个第二序列;

[0253] 其中,第一序列,用于表示在第一时段至少一个运行参数的状态变化的序列。

[0254] 处理模块42,用于基于信息分类规则,得到第一分类规则和第二分类规则;基于第一分类规则和第二分类规则,对第一数据中的运行参数进行处理,得到至少一个第一序列;其中,第一分类规则,用于表示第一数据中的至少一个运行参数的类别划分规则;第二分类规则,用于表示第一数据中的至少一个运行参数的参数值的区间划分规则。

[0255] 处理模块42,用于获取训练完成的循环神经网络;其中,循环神经网络,用于预测数据库的第二时刻和/或第二时段的至少一个运行状态参数值;

[0256] 基于训练完成的循环神经网络和至少一个第二序列,对第二数据进行处理,得到第一信息。

[0257] 处理模块42,用于基于第一数据,对循环神经网络进行训练,得到训练完成的循环神经网络。

[0258] 处理模块42,用于基于训练完成的循环神经网络,对第二数据进行处理,得到至少一个第三序列;基于至少一个第三序列和至少一个第二序列,得到第一信息;其中,至少一个第三序列,包括数据库未来时刻的至少一个运行参数信息值。

[0259] 处理模块42,用于对至少一个第二序列进行聚类,得到至少一个类簇;

[0260] 基于至少一个类簇,对至少一个第三序列进行处理,得到第一信息。

[0261] 处理模块42,用于将第二数据输入至训练完成的循环神经网络,得到第三数据;

[0262] 基于信息分类规则,对第三数据进行处理,得到至少一个第三序列。

[0263] 基于至少一个第二序列,对第二数据进行处理,得到第一信息。

[0264] 在实际应用中,上述获取模块41、处理模块42均可由位于电子设备中的处理器实现,上述处理器为ASIC、DSP、DSPD、PLD、FPGA、CPU、控制器、微控制器、微处理器中的至少一种。

[0265] 本申请实施例提供的信息处理设备4,首先获取表征数据库历史运行过程中的第一时刻的至少一种运行状态信息,还要获取表示第一时刻之前预设长度时间段的至少一个运行参数,然后获取表示数据库当前时刻的至少一个运行参数的第二数据,再基于第一数据和第二数据,得到表示数据库的未来时刻的运行状态信息。由此,本申请实施例所提供的信息处理设备4,可以根据数据库历史运行过程当中的运行状态信息和至少一个运行参数,结合当前时刻的至少一个运行参数的信息,预测数据库未来时刻和/或未来时段的运行状态信息,从而解决了相关技术中在数据库尚未出现异常时,无法预测数据库在未来的运行中是否会出现异常的问题。

[0266] 基于前述实施例,本申请实施例提供了一种计算机可读存储介质,该计算机可读存储介质存储有一个或多个程序,一个或多个程序可以被多个处理器执行,以实现如前所述的任一信息处理方法的步骤。

[0267] 需要说明的是,上述计算机可读存储介质可以是只读存储器(Read Only Memory, ROM)、可编程只读存储器(Programmable Read-Only Memory, PROM)、可擦除可编程只读存储器(Erasable Programmable Read-Only Memory, EPROM)、电可擦除可编程只读存储器(Electrically Erasable Programmable Read-Only Memory, EEPROM)、磁性随机存取存储器(Ferromagnetic Random Access Memory, FRAM)、快闪存储器(Flash Memory)、磁表面存储器、光盘、或只读光盘(Compact Disc Read-Only Memory, CD-ROM)等存储器;也可以是包括上述存储器之一或任意组合的各种电子设备,如移动电话、计算机、平板设备、个人数字助理等。

[0268] 上文对各个实施例的描述倾向于强调各个实施例之间的不同之处,其相同或相似之处可以互相参考,为了简洁,本文不再赘述。

[0269] 本申请所提供的各方法实施例中所揭露的方法,在不冲突的情况下可以任意组合,得到新的方法实施例。

[0270] 本申请所提供的各产品实施例中所揭露的特征,在不冲突的情况下可以任意组合,得到新的产品实施例。

[0271] 本申请所提供的各方法或设备实施例中所揭露的特征,在不冲突的情况下可以任意组合,得到新的方法实施例或设备实施例。

[0272] 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过程、方法、物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,并不排除在包括该要素的过程、方法、物品或者装置中还存在另外的相同要素。

[0273] 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。

[0274] 通过以上的实施方式的描述,本领域的技术人员可以清楚地了解到上述实施例方

法可借助软件加必需的通用硬件平台的方式来实现,当然也可以通过硬件,但很多情况下前者是更佳的实施方式。基于这样的理解,本申请的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质(如ROM/RAM、磁碟、光盘)中,包括若干指令用以使得一台终端设备(可以是手机,计算机,服务器,空调器,或者网络设备等)执行本申请各个实施例所描述的方法。

[0275] 本申请是参照根据本申请实施例的方法、设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器,使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

[0276] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中,使得存储在该计算机可读存储器中的指令产生包括指令装置的制品,该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

[0277] 这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上,使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理,从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

[0278] 以上仅为本申请的优选实施例,并非因此限制本申请的专利范围,凡是利用本申请说明书及附图内容所作的等效结构或等效流程变换,或直接或间接运用在其他相关的技术领域,均同理包括在本申请的专利保护范围内。

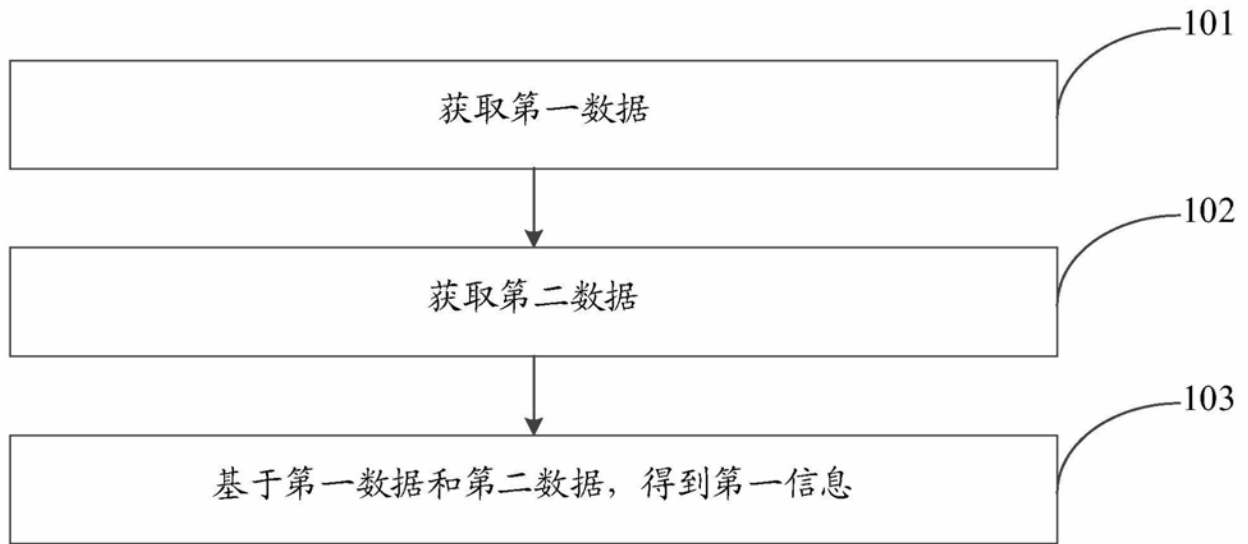


图1

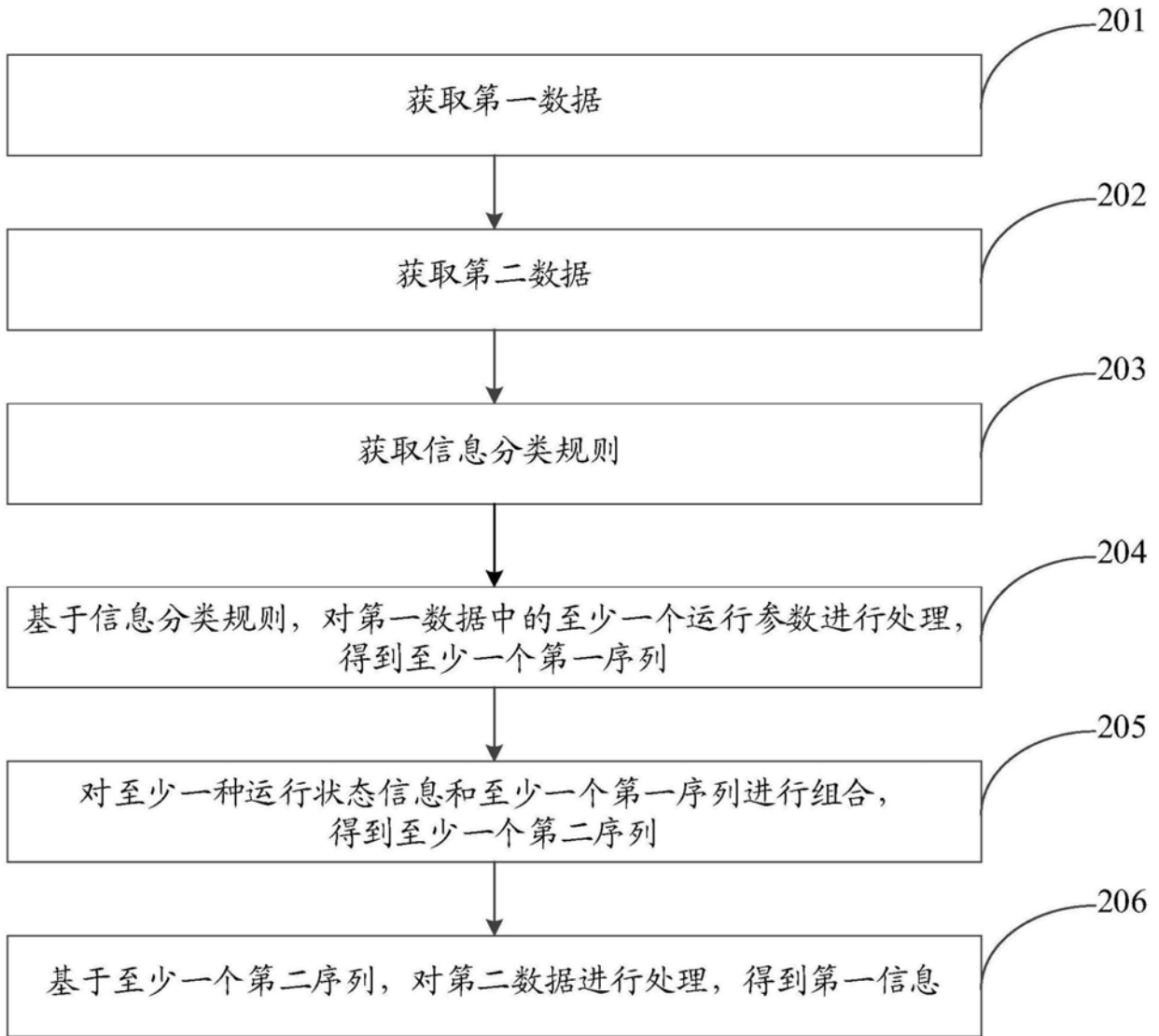


图2

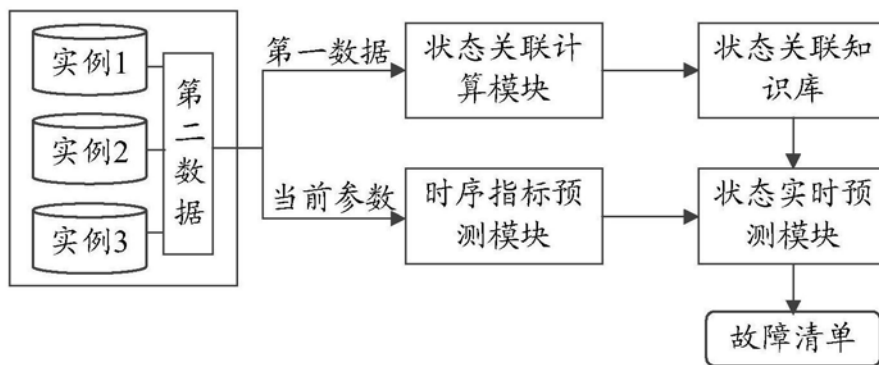


图3

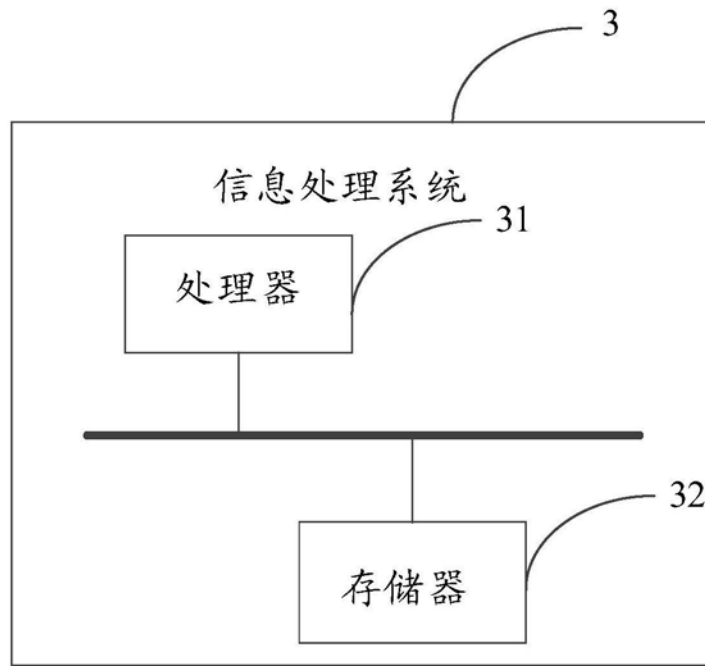


图4

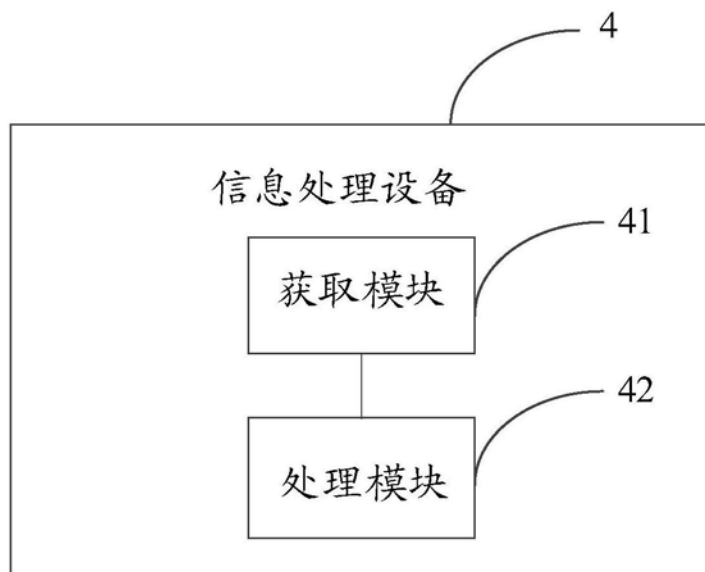


图5