



(12) 发明专利申请

(10) 申请公布号 CN 112052151 A

(43) 申请公布日 2020. 12. 08

(21) 申请号 202011072717.9

G06K 9/62 (2006.01)

(22) 申请日 2020.10.09

G06N 3/04 (2006.01)

G06N 3/08 (2006.01)

(71) 申请人 腾讯科技(深圳)有限公司

地址 518057 广东省深圳市南山区高新区
科技中一路腾讯大厦35层

(72) 发明人 刘志煌 胡林红 罗朝亮 武睿彪
李冠灿

(74) 专利代理机构 广州三环专利商标代理有限
公司 44202

代理人 熊永强 贾允

(51) Int. Cl.

G06F 11/34 (2006.01)

G06F 16/17 (2019.01)

G06F 16/2458 (2019.01)

G06F 16/26 (2019.01)

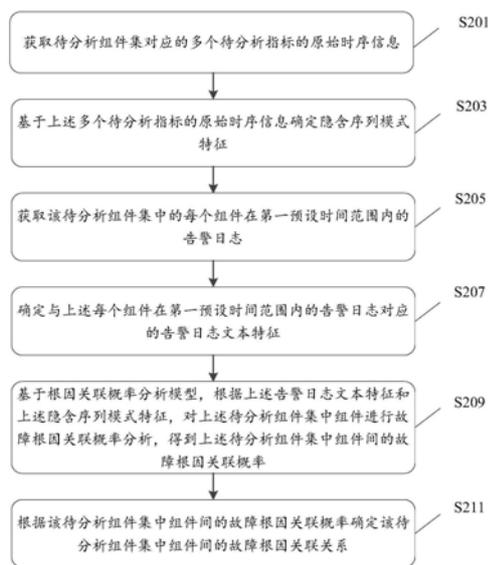
权利要求书2页 说明书17页 附图8页

(54) 发明名称

故障根因分析方法、装置、设备及存储介质

(57) 摘要

本申请公开了一种故障根因分析方法、装置、设备及存储介质,所述方法包括获取待分析组件集对应的多个待分析指标的原始时序信息;基于原始时序信息确定隐含序列模式特征;获取待分析组件集中的每个组件在第一预设时间范围内的告警日志;确定与每个组件在第一预设时间范围内的告警日志对应的告警日志文本特征;基于根因关联概率分析模型,根据告警日志文本特征和隐含序列模式特征,对待分析组件集中组件进行故障根因关联概率分析,得到待分析组件集中组件间的故障根因关联概率;根据故障根因关联概率确定组件间的故障根因关联关系。利用本申请提供的技术方案能够在故障检测中高效准确地确定组件间的故障根因关联关系,提升故障根因分析的可靠性。



1. 一种故障根因分析方法,其特征在于,所述方法包括:

获取待分析组件集对应的多个待分析指标的原始时序信息,所述多个待分析指标包括所述待分析组件集中的每个组件对应的待分析指标;

基于所述多个待分析指标的原始时序信息确定隐含序列模式特征;

获取所述待分析组件集中的每个组件在第一预设时间范围内的告警日志;

确定与所述每个组件在第一预设时间范围内的告警日志对应的告警日志文本特征;

基于根因关联概率分析模型,根据所述告警日志文本特征和所述隐含序列模式特征,对所述待分析组件集中组件进行故障根因关联概率分析,得到所述待分析组件集中组件间的故障根因关联概率;

根据所述待分析组件集中组件间的故障根因关联概率确定所述待分析组件集中组件间的故障根因关联关系。

2. 根据权利要求1所述的方法,其特征在于,所述根因关联概率分析模型包括相关性挖掘模块、特征融合层、前馈层和分类层;

所述基于根因关联概率分析模型,根据所述告警日志文本特征和所述隐含序列模式特征,对所述待分析组件集中组件进行故障根因关联概率分析,得到所述待分析组件集中组件间的故障根因关联概率包括:

基于所述相关性挖掘模块对所述告警日志文本特征进行相关性挖掘,得到告警日志相关性特征;

基于所述特征融合层对所述告警日志相关性特征和所述隐含序列模式特征进行特征融合,得到目标融合特征;

基于所述前馈层对所述目标融合特征进行特征处理,得到处理后的目标融合特征;

基于所述分类层对所述处理后的目标融合特征进行故障根因关联概率计算,得到所述待分析组件集中组件间的故障根因关联概率。

3. 根据权利要求1所述的方法,其特征在于,所述基于所述多个待分析指标的原始时序信息确定隐含序列模式特征包括:

根据所述多个待分析指标的原始时序信息确定第二预设时间范围内的指标时序升降序列;

根据所述指标时序升降序列进行序列模式挖掘,得到隐含序列模式;

对所述隐含序列模式进行特征编码,得到所述隐含序列模式特征。

4. 根据权利要求3所述的方法,其特征在于,所述根据所述多个待分析指标的原始时序信息确定第二预设时间范围内的指标时序升降序列包括:

根据所述多个待分析指标的原始时序信息确定所述多个待分析指标的时序升降信息;

根据所述多个待分析指标的时序升降信息构造所述第二预设时间范围内的指标时序升降序列。

5. 根据权利要求1所述的方法,其特征在于,所述方法还包括:

获取多个标注有组件间的故障根因关联概率的样本组件集对应的样本隐含序列模式特征,以及对应的样本告警日志文本特征;

基于所述多个标注有组件间的故障根因关联概率的样本组件集对应的样本隐含序列模式特征,以及对应的样本告警日志文本特征,对预设神经网络模型进行故障根因关联概

率分析的训练,在故障根因关联概率分析的训练中调整所述预设神经网络模型的模型参数直至所述预设神经网络模型满足预设收敛条件,得到所述故障根因关联概率分析模型。

6. 根据权利要求5所述的方法,其特征在于,所述获取多个标注有组件间的故障根因关联概率的样本组件集对应的样本告警日志文本特征包括:

分别获取所述样本组件集中的每个组件在第三预设时间范围内的样本告警日志;
分别对所述样本告警日志进行文本向量化,得到对应的样本告警日志文本特征。

7. 根据权利要求1所述的方法,其特征在于,所述根据所述待分析组件集中组件间的故障根因关联概率确定所述待分析组件集中组件间的故障根因关联关系包括:

当所述故障根因关联概率满足预设条件时,确定所述待分析组件集中组件间的故障根因关联关系为存在故障根因关联。

8. 一种故障根因分析装置,其特征在于,所述装置包括:

原始时序信息获取模块,用于获取待分析组件集对应的多个待分析指标的原始时序信息,所述多个待分析指标包括所述待分析组件集中的每个组件对应的待分析指标;

隐含序列模式特征确定模块,用于基于所述多个待分析指标的原始时序信息确定隐含序列模式特征;

告警日志获取模块,用于获取所述待分析组件集中的每个组件在第一预设时间范围内的告警日志;

文本特征确定模块,用于确定与所述每个组件在第一预设时间范围内的告警日志对应的告警日志文本特征;

故障根因关联概率分析模块,用于基于根因关联概率分析模型,根据所述告警日志文本特征和所述隐含序列模式特征,对所述待分析组件集中组件进行故障根因关联概率分析,得到所述待分析组件集中组件间的故障根因关联概率;

故障根因关联关系确定模块,用于根据所述待分析组件集中组件间的故障根因关联概率确定所述待分析组件集中组件间的故障根因关联关系。

9. 一种故障根因分析设备,其特征在于,所述设备包括处理器和存储器,所述存储器中存储有至少一条指令或至少一段程序,所述至少一条指令或所述至少一段程序由所述处理器加载并执行以实现如权利要求1至7任一所述的故障根因分析方法。

10. 一种计算机可读存储介质,其特征在于,所述存储介质中存储有至少一条指令或至少一段程序,所述至少一条指令或所述至少一段程序由处理器加载并执行以实现如权利要求1至7任一所述的故障根因分析方法。

故障根因分析方法、装置、设备及存储介质

技术领域

[0001] 本申请涉及运维管理技术领域,具体涉及一种故障根因分析方法、装置、设备及存储介质。

背景技术

[0002] 随着数字化转型的不断推进,各种系统的数据指标和调用关系变得越来越复杂,一个系统往往由大量的服务器等组件构成,一旦发生故障可能会带来巨大的损失,因此除了需要迅速检测之外,还需要进行故障根因分析,从而避免以后再次发生类似的故障,减少故障带来的损失。

[0003] 现有技术在进行故障根因分析时,往往需要人工指定规则或积累经验,构建决策树,或是建立知识图谱,灵活性低,对人工的依赖性强,效率低且难免出错,在规则等需要更新时耗费的时间和人力资源较多,需要提供更加可靠高效的方案。

发明内容

[0004] 为了解决现有技术的问题,本申请提供了一种故障根因分析方法、装置、设备及存储介质。所述技术方案如下:

[0005] 本申请一方面提供了一种故障根因分析方法,所述方法包括:

[0006] 获取待分析组件集对应的多个待分析指标的原始时序信息,所述多个待分析指标包括所述待分析组件集中的每个组件对应的待分析指标;

[0007] 基于所述多个待分析指标的原始时序信息确定隐含序列模式特征;

[0008] 获取所述待分析组件集中的每个组件在第一预设时间范围内的告警日志;

[0009] 确定与所述每个组件在第一预设时间范围内的告警日志对应的告警日志文本特征;

[0010] 基于根因关联概率分析模型,根据所述告警日志文本特征和所述隐含序列模式特征,对所述待分析组件集中组件进行故障根因关联概率分析,得到所述待分析组件集中组件间的故障根因关联概率;

[0011] 根据所述待分析组件集中组件间的故障根因关联概率确定所述待分析组件集中组件间的故障根因关联关系。

[0012] 本申请另一方面提供了一种故障根因分析装置,所述装置包括:

[0013] 原始时序信息获取模块,用于获取待分析组件集对应的多个待分析指标的原始时序信息,所述多个待分析指标包括所述待分析组件集中的每个组件对应的待分析指标;

[0014] 隐含序列模式特征确定模块,用于基于所述多个待分析指标的原始时序信息确定隐含序列模式特征;

[0015] 告警日志获取模块,用于获取所述待分析组件集中的每个组件在第一预设时间范围内的告警日志;

[0016] 文本特征确定模块,用于确定与所述每个组件在第一预设时间范围内的告警日志

对应的告警日志文本特征；

[0017] 故障根因关联概率分析模块,用于基于根因关联概率分析模型,根据所述告警日志文本特征和所述隐含序列模式特征,对所述待分析组件集中组件进行故障根因关联概率分析,得到所述待分析组件集中组件间的故障根因关联概率；

[0018] 故障根因关联关系确定模块,用于根据所述待分析组件集中组件间的故障根因关联概率确定所述待分析组件集中组件间的故障根因关联关系。

[0019] 本申请另一方面提供了一种故障根因分析设备,所述设备包括处理器和存储器,所述存储器中存储有至少一条指令或至少一段程序,所述至少一条指令或所述至少一段程序由上述处理器加载并执行以实现如上述的故障根因分析方法。

[0020] 本申请另一方面提供了一种计算机可读存储介质,所述存储介质中存储有至少一条指令或至少一段程序,所述至少一条指令或所述至少一段程序由处理器加载并执行以实现如上述的故障根因分析方法。

[0021] 本申请提供的故障根因分析方法、装置、设备及存储介质,具有如下技术效果:

[0022] 本申请通过获取待分析组件集对应的多个待分析指标的原始时序信息,可以确定隐含序列模式特征;且获取待分析组件集中的每个组件在第一预设时间范围内的告警日志来确定对应的告警日志文本特征,适应动态运维变化的需求;然后,基于根因关联概率分析模型,根据告警日志文本特征和隐含序列模式特征,对待分析组件集中组件进行故障根因关联概率分析,得到待分析组件集中组件间的故障根因关联概率,能够快速准确地得到待分析组件集中组件间的故障根因关联概率,最后,根据待分析组件集中组件间的故障根因关联概率确定待分析组件集中组件间的故障根因关联关系。利用本说明书实施例提供的技术方案可以快速准确地确定组件间的故障根因关联关系,进而提升故障根因分析的可靠性。

[0023] 本申请的附加方面和优点将在下面的描述中部分给出,部分将从下面的描述中变得明显,或通过本申请的实践了解到。

附图说明

[0024] 为了更清楚地说明本申请实施例或现有技术中的技术方案和优点,下面将对实施例或现有技术描述中所需要使用的附图作简单的介绍,显而易见地,下面描述中的附图仅仅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它附图。

[0025] 图1是本申请实施例提供的一种应用环境的示意图；

[0026] 图2是本申请实施例提供的一种故障根因分析方法的流程示意图；

[0027] 图3是本申请实施例提供的另一种故障根因分析方法的流程示意图；

[0028] 图4是本申请实施例提供的另一种故障根因分析方法的流程示意图；

[0029] 图5是本申请实施例提供的一种根因关联概率分析模型的结构示意图；

[0030] 图6是本申请实施例提供的另一种故障根因分析方法的流程示意图；

[0031] 图7是本申请实施例提供的另一种根因关联概率分析模型的结构示意图；

[0032] 图8是本申请实施例提供的另一种故障根因分析方法的流程示意图；

[0033] 图9是本申请实施例提供的一种故障根因分析装置示意图；

[0034] 图10是本申请实施例提供的一种用于实现故障根因分析方法的服务器的硬件结构框图。

具体实施方式

[0035] 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,所描述的实施例仅仅是本申请一部分实施例,而不是全部的实施例。基于本申请中的实施例,本领域普通技术人员在没有做出创造性劳动的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。所述实施例的示例在附图中示出,其中自始至终相同或类似的标号表示相同或类似的元件或具有相同或类似功能的元件。

[0036] 需要说明的是,本申请的说明书和权利要求书及上述附图中的术语“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。应该理解这样使用的数据在适当情况下可以互换,以便这里描述的本申请的实施例能够以除了在这里图示或描述的那些以外的顺序实施。此外,术语“包括”和“具有”以及他们的任何变形,意图在于覆盖不排他的包含,例如,包含了一系列步骤或单元的过程、方法、系统、产品或服务不必限于清楚地列出的那些步骤或单元,而是可包括没有清楚地列出的或对于这些过程、方法、产品或服务固有的其它步骤或单元。

[0037] 请参阅图1,图1是本申请提供的一种应用环境示意图,如图1所示,该应用环境可以包括根因分析服务器01和多个服务组件02。

[0038] 在本说明书实施例中,根因分析服务器01可以用于结合多个服务组件02的数据进行故障根因分析,可选的,根因分析服务器01可以是独立的物理服务器,也可以是多个物理服务器构成的服务器集群或者分布式系统,还可以是提供云服务、云数据库、云计算、云函数、云存储、网络服务、云通信、中间件服务、域名服务、安全服务、CDN(Content Delivery Network内容分发网络)、以及大数据和人工智能平台等基础云计算服务的云服务器。

[0039] 在本说明书实施例中,多个服务组件02可以生成运行数据及告警日志等,以使根因分析服务器01能够获取需要的数据来实现故障根因分析,在一个实施例中,上述的多个服务组件02可以包括用于实现不同功能的服务器,可以是独立的物理服务器,也可以是多个物理服务器构成的服务器集群或者分布式系统,还可以是提供云服务、云数据库、云计算、云函数、云存储、网络服务、云通信、中间件服务、域名服务、安全服务、CDN(Content Delivery Network内容分发网络)、以及大数据和人工智能平台等基础云计算服务的云服务器。在实际应用中,服务组件02还可以包括但不限于智能手机、平板电脑、笔记本电脑、台式计算机、智能音箱、智能手表等终端设备,以及网络设备和防火墙等。

[0040] 在本说明书实施例中,上述的根因分析服务器01以及多个服务组件02可以通过有线或无线通信方式进行直接或间接地连接,本申请在此不做限制。

[0041] 图2是本申请实施例提供的一种故障根因分析方法的流程图,本说明书提供了如实施例或流程图所述的方法操作步骤,但基于常规或者无创造性的劳动可以包括更多或者更少的操作步骤。实施例中列举的步骤顺序仅仅为众多步骤执行顺序中的一种方式,不代表唯一的执行顺序。在实际中的系统或服务产品执行时,可以按照实施例或者附图所示的方法顺序执行或者并行执行(例如并行处理器或者多线程处理的环境)。具体的如图2所示,所述方法可以包括:

[0042] S201:获取待分析组件集对应的多个待分析指标的原始时序信息。

[0043] 在本说明书实施例中,该待分析组件集包括至少两个组件。具体地,该待分析组件集中的组件可以结合实际的故障根因分析需求进行设定,在一个具体的实施例中,该待分析组件集可以包括在一次异常事件中发生故障的组件,以及可能与该发生故障的组件相关的至少一个组件。

[0044] 在本说明书实施例中,上述的故障根因分析可以包括分析预设几个组件是否存在故障关联,避免再次发生类似故障;由于在实际应用中,例如系统中的一个服务器发生故障了,那么与这个服务器相关的其他组件(例如防火墙、网络设备和传输设备等)有可能也发生了故障,但也可能没有发生故障,这些组件中可以包括与这个故障的服务器存在故障根因关联的组件(例如网络设备的某些参数短暂异常是该服务器故障的因素之一,但该网络设备仍可正常运行),在本说明书实施例中,可以结合实际的故障根因分析需求,将预设的几个组件作为待分析组件集,后续可以基于根因关联概率分析模型进行故障根因关联概率分析,以确定待分析组件集中组件间的故障根因关联关系,有利于运维人员后续进行相应的维护,避免再次发生类似故障。

[0045] 具体地,组件可以包括但不限于终端设备、用于实现不同功能的服务器、网络设备和防火墙等;指标可以用于表征对应的组件的相关运行信息,具体地,指标可以包括但不限于平均响应时间、平均吞吐率、请求数、错误率、健康度和处理耗时。

[0046] 在本说明书实施例中,上述的多个待分析指标包括该待分析组件集中的每个组件对应的待分析指标。由于每个组件可能对应多个指标,可以结合实际故障根因分析需求获取每个组件对应的全部指标中的某几个指标作为该组件对应的待分析指标。例如,待分析组件集包括A组件、B组件和C组件,可以获取A组件对应的全部指标中的3个指标作为A组件对应的待分析指标,获取B组件对应的全部指标中的5个指标作为B组件对应的待分析指标,获取C组件对应的全部指标中的2个指标作为C组件对应的待分析指标,此时可将这10个指标作为上述的待分析组件集对应的多个待分析指标。

[0047] 在本说明书实施例中,每个待分析指标的原始时序信息可以表征该待分析指标的值随时间的变化关系,在一个实施例中,该原始时序信息可以包括随时间连续变化的二维曲线,或,多个随时间离散变化的点值。例如,当待分析指标包括A组件的平均吞吐率时,该待分析指标的原始时序信息可以为随时间变化的二维曲线,横坐标为时间,纵坐标为平均吞吐率的值,获取利用待分析指标的原始时序信息可以获取该待分析指标在每个时刻的值以及变化趋势。在实际应用中,可以结合实际故障根因分析的需求获取待分析指标在任意时间的原始时序信息,较为灵活。

[0048] S203:基于上述多个待分析指标的原始时序信息确定隐含序列模式特征。

[0049] 在一个具体的实施例中,如图3所示,基于上述多个待分析指标的原始时序信息确定隐含序列模式特征可以包括:

[0050] S301:根据上述多个待分析指标的原始时序信息确定第二预设时间范围内的指标时序升降序列。

[0051] 在本说明书实施例中,一个指标时序升降序列可以包括多个指标变化标识符,该指标变化标识符可以表征对应的待分析指标的变化情况,在一个具体的实施例中,当上述的多个待分析指标包括a指标、b指标和c指标,上述的指标变化标识符可以包括,例如,a增,

a减,b增,b减,c增,c减。

[0052] 在一个实施例中,当原始时序信息包括随时间连续变化的二维曲线时,可以通过确定曲线的变化节点来确定该原始时序信息的增减性的变化(例如原本在该变化节点之前曲线变化趋势为递增,在该变化节点之后曲线变化趋势变为递减;或;原本在该变化节点之前曲线变化趋势为递减,在该变化节点之后曲线变化趋势变为递增)。具体地,上述的第二预设时间范围可以包括多个预设连续时间段,且上述预设连续时间段可以结合实际故障根因分析需求进行确定。上述的根据上述多个待分析指标的原始时序信息确定第二预设时间范围内的指标时序升降序列可以包括:

[0053] 分别基于每个预设连续时间段内上述多个待分析指标的原始时序信息的变化节点出现次序确定对应的指标时序升降序列,将全部预设连续时间段对应的指标时序升降序列作为上述的第二预设时间范围内的指标时序升降序列。

[0054] 在一个具体的实施例中,第二预设时间范围可以包括7月9日的20~23时、7月10日的20~23时和7月11日的20~23时3个预设连续时间段,上述的待分析指标包括a指标、b指标c指标和d指标,在7月9日的20~23时这个预设连续时间段内,b指标对应的曲线先出现了变化节点,b指标对应的曲线在该变化节点后变为了b增,然后,c指标对应的曲线出现了变化节点,c指标对应的曲线在该变化节点后变为了c增,然后,a指标对应的曲线出现了变化节点,a指标对应的曲线在该变化节点后变为了a减,然后,d指标对应的曲线出现了变化节点,d指标对应的曲线在该变化节点后变为了d增,则此时确定出的对应的指标时序升降序列为“b增-c增-a减-d增”,该指标时序升降序列包括指标变化标识符b增、c增、a减和d增;同理,可以确定另外2个预设连续时间段对应的指标时序升降序列。

[0055] 通过分别基于每个预设连续时间段内上述多个待分析指标的原始时序信息的变化节点出现次序确定对应的指标时序升降序列,有利于确定多个指标的变化间是否存在潜在的因果关系,从而有利于后续按需进行故障根因分析,提升故障根因分析的可靠性和全面性。

[0056] 在前一个实施例中,可以通过确定曲线的变化节点来确定该原始时序信息的增减性的变化,分别基于每个预设连续时间段内上述多个待分析指标的原始时序信息的变化节点出现次序来确定对应的指标时序升降序列。在本说明书提供的另一个实施例中,还可以通过基于原始时序信息设置多个时间间隔,通过将—个时间间隔内的指标的—值与对应的前—个时间间隔内的该指标的—值比较,来确定待分析指标在每个时间间隔内指标的—值的增减,进而确定第二预设时间范围内的指标时序升降序列。在该实施例中,具体地,如图4所示,根据上述多个待分析指标的原始时序信息确定第二预设时间范围内的指标时序升降序列可以包括:

[0057] S401:根据上述多个待分析指标的原始时序信息确定上述多个待分析指标的时序升降信息。

[0058] 具体地,上述的根据上述多个待分析指标的原始时序信息确定上述多个待分析指标的时序升降信息可以包括:

[0059] 1) 设置多个时间节点,将每两个相邻时间节点之间的时间区间作为—个时间间隔;

[0060] 2) 分别根据每个待分析指标的原始时序信息确定该待分析指标的—值在每个时间

间隔的增减信息；

[0061] 在实际应用中,根据每个待分析指标的原始时序信息确定该待分析指标的在每个时间间隔的增减信息,可以通过将该待分析指标在一个时间间隔内的值与对应的前一个时间间隔内的值比较,来确定该待分析指标的在该时间间隔的增减信息。

[0062] 3) 根据该待分析指标的在每个时间间隔的增减信息进行时序升降标记,得到该待分析指标的时序升降信息,整合上述多个待分析指标的时序升降信息。

[0063] 例如,当确定a指标在1-2时的值相比a指标在0-1时的值为增加,则可以标记a指标在1-2时对应的指标变化标识符为a增。

[0064] 在一个具体的实施例中,可以每隔1小时设置一个时间节点,上述的多个待分析指标包括a指标、b指标、c指标、d指标、e指标和f指标,以2020年7月1日的0-12时为例,上述多个待分析指标的时序升降信息可以利用下述的

[0065] 表1来表示:

[0066]

时间间隔	a指标	b指标	c指标	d指标	e指标	f指标
0-1时	a减	b增	c增	d增	e减	f增
1-2时	a增	b减	c增	d减	e增	f减
2-3时	a减	b增	c减	d增	e增	f减
3-4时	a减	b减	c增	d减	e减	f减
4-5时	a增	b增	c增	d增	e增	f增
5-6时	a增	b减	c增	d减	e减	f增
6-7时	a增	b增	c减	d减	e增	f减
7-8时	a增	b减	c减	d增	e增	f增
8-9时	a增	b增	c增	d减	e增	f减
9-10时	a减	b增	c增	d减	e减	f增
10-11时	a减	b减	c增	d增	e增	f增
11-12时	a增	b减	c减	d减	e减	f增

[0067] 表1

[0068] 其他日期的上述多个待分析指标的时序升降信息可以与表1的形式类似,在此不再赘述。

[0069] S403:根据上述多个待分析指标的时序升降信息构造第二预设时间范围内的指标时序升降序列。

[0070] 在实际应用中,上述的第二预设时间范围可以结合实际故障根因分析需求进行设定,在一个实施例中,上述的第二预设时间范围可以取不同日期的相同时间段(时间间隔),在另一个实施例中,上述的第二预设时间范围还可以取同一天的不同时间段(时间间隔),本申请并不以此为限。以上述的第二预设时间范围取不同日期的相同时间段为例,可以根据已经生成的多个待分析指标的时序升降信息,分别构造2020年7月1日的8-9时这个时间间隔对应的指标时序升降序列,以及2020年7月2日的8-9时这个时间间隔对应的指标时序升降序列,如表2所示:

[0071]

日期	时间间隔	指标时序升降序列
20200701	8-9时	a增-b增-c增-d减-e增-f减

20200702	8-9时	a减-b增-c增-d减-e减-f减
----------	------	-------------------

[0072] 表2

[0073] 以“a增-b增-c增-d减-e增-f减”这个指标时序升降序列为例,可以理解为此时a增伴随着b增伴随着c增伴随着d减伴随着e增伴随着f减,通过根据上述多个待分析指标的原始时序信息确定上述多个待分析指标的时序升降信息,再根据上述多个待分析指标的时序升降信息构造第二预设时间范围内的指标时序升降序列,有利于获取大量的指标变化信息以便确定多个指标的变化间是否存在潜在的关联关系,从而有利于后续按需进行故障根因分析,提升故障根因分析的可靠性。

[0074] S303:根据上述指标时序升降序列进行序列模式挖掘,得到隐含序列模式。

[0075] 由于各指标的变化可能存在潜在的因果关系或关联关系,有必要根据上述指标时序升降序列进行序列模式挖掘,在本说明书实施例中,可以根据上述指标时序升降序列,利用prefixspan算法(Prefix-Projected Pattern Growth,前缀投影的序列模式挖掘)进行序列模式挖掘,得到隐含序列模式。具体地,根据上述指标时序升降序列进行序列模式挖掘,得到隐含序列模式可以包括如下步骤:

[0076] 1) 确定上述的指标时序升降序列中每个指标变化标识符的频数;

[0077] 具体地,上述的频数可以表征该指标变化标识符在全部指标时序升降序列中的出现次数。

[0078] 以上述的表2为例,此时共有2个指标时序升降序列,即“a增-b增-c增-d减-e增-f减”和“a减-b增-c增-d减-e减-f减”,确定的上述的指标时序升降序列中每个指标变化标识符的频数如表3所示:

[0079] 指标变化标识符	a增	a减	b增	c增	d减	e增	e减	f减
频数	1	1	2	2	2	1	1	2

[0080] 表3

[0081] 2) 基于上述指标变化标识符的频数确定满足预设最小支持度阈值的指标变化标识符,分别将上述满足预设最小支持度阈值的指标变化标识符作为一项前缀,并确定对应的后缀。

[0082] 在本说明书实施例中,该预设最小支持度阈值可以结合实际应用需求进行设定,在一个实施例中,该预设最小支持度阈值可以结合下述公式确定:

$$[0083] \quad \min_sup = a \times n$$

[0084] 其中,min_sup表示上述的预设最小支持度阈值,n表示第二预设时间范围包括的天数(日期数),a表示最小支持率,该最小支持率可以结合实际应用需求确定,例如,结合指标时序升降序列数量进行调整,随着指标时序升降序列数量,可以调低最小支持率。该预设最小支持度阈值可以表征对于数据出现的频繁程度的要求,例如,该预设最小支持度阈值为0.5,则当目标数据在所有数据中的出现频率高于0.5时满足该预设最小支持度阈值,假如共有10个指标时序升降序列,当目标元素在多于5个指标时序升降序列中出现时,确定该目标元素满足该预设最小支持度阈值。

[0085] 请参照表4,当该预设最小支持度阈值为0.5时,步骤2) 确定出的一项前缀和对应的后缀如表4所示:

一项前缀	对应后缀
b 增	c 增-d 减-e 增-f 减
	c 增-d 减-e 减-f 减
c 增	d 减-e 增-f 减
	d 减-e 减-f 减
d 减	e 增-f 减
	e 减-f 减
f 减	无

[0087] 表4

[0088] 3) 分别确定各一项前缀对应的后缀中满足该预设最小支持度阈值的单项,将满足该预设最小支持度阈值的单项与对应的一项前缀合并,得到二项前缀,继续确定各二项前缀对应的后缀。

[0089] 请参照表5,当该预设最小支持度阈值为0.5时,步骤3) 确定出的各二项前缀和对应的后缀如表5所示:

二项前缀	对应后缀
b 增- c 增	d 减-e 增-f 减
	d 减-e 减-f 减
c 增- d 减	e 增-f 减
	e 减-f 减
d 减- f 减	无

[0091] 表5

[0092] 4) 以此类推,分别确定各i项前缀对应的后缀中满足该预设最小支持度阈值的单项,将该满足该预设最小支持度阈值的单项与对应的i项前缀合并,得到(i+1)项前缀,并确定与(i+1)项前缀对应的后缀(i为大于1的整数);

[0093] 重复执行步骤4),直到挖掘出最长前缀序列,将该最长前缀序列作为上述的隐含序列模式。

[0094] 请参照表6和表7,当该预设最小支持度阈值为0.5时,确定出的各三项前缀和对应的后缀如表6所示,各四项前缀和对应的后缀如表7所示:

[0095]	三项前缀	对应后缀
	b增-c增-d减	e增-f减 e减-f减
	c增-d减-f减	无

[0096] 表6

[0097]	四项前缀	对应后缀
	b增-c增-d减-f减	无

[0098] 表7

[0099] 此时挖掘出的最长前缀序列即“b增-c增-d减-f减”，即根据如表2所示的指标时序升降序列进行序列模式挖掘得到的隐含序列模式即“b增-c增-d减-f减”。通过根据上述多个待分析指标的原始时序信息确定第二预设时间范围内的指标时序升降序列，根据上述指标时序升降序列进行序列模式挖掘，得到隐含序列模式，该隐含序列模式可以为上述多个待分析指标变化隐含的规律，可能是几个指标变化的关联关系或因果关系，后续可以对隐含序列模式进行特征编码，结合各组件的告警日志进行故障根因分析，有利于提升故障根因分析的可靠性；由于各指标数据会随时间不断地更新变化，该隐含序列模式也会不断变化，某几个指标在过去一段时间没有关联，但之后可能会存在关联，因此可以根据需求调整上述的第二预设时间范围，来实时挖掘最新的隐含序列模式，灵活性强，提升故障根因分析的时效性。

[0100] S307:对该隐含序列模式进行特征编码，得到隐含序列模式特征。

[0101] 在本说明书实施例中，可以对该隐含序列模式进行One-Hot特征编码(独热编码)，得到隐含序列模式特征。

[0102] S205:获取该待分析组件集中的每个组件在第一预设时间范围内的告警日志。

[0103] 具体地，该第一预设时间范围可以结合实际故障根因分析需求进行设定；在一个具体的实施例中，该第一预设时间范围可以包括故障发生时刻之前的一个小时至故障发生时刻之后的一个小时。例如该待分析组件集包括A组件、B组件和C组件，A组件在第一预设时间范围内生成了4个告警日志，B组件在第一预设时间范围内生成了3个告警日志，C组件在第一预设时间范围内生成了3个告警日志，可以获取这10个告警日志，后续分别确定每个告警日志对应的告警日志文本特征。

[0104] 由于告警日志属于半结构化数据，特点是实时而且数据丰富，通过获取该待分析组件集中的每个组件在第一预设时间范围内的告警日志，后续可以结合隐含序列模式进行故障根因分析，有利于提升故障根因分析的可靠性。

[0105] S207:确定与上述每个组件在第一预设时间范围内的告警日志对应的告警日志文本特征。

[0106] 在本说明书实施例中，确定与上述每个组件在第一预设时间范围内的告警日志对应的告警日志文本特征可以包括：

[0107] 分别对每个告警日志进行文本向量化，得到对应的告警日志文本特征。

[0108] 在一个具体的实施例中，分别对每个告警日志进行文本向量化，得到对应的告警

日志文本特征可以包括：

[0109] 1) 基于预设词向量模型获取该告警日志中每个词对应的词向量；

[0110] 在实际应用中,该预设词向量模型可以包括Word2vec词向量模型。需要说明的是,当上述告警日志的文本为预设文本类型,例如,中文时,在基于预设词向量模型获取该告警日志中每个词对应的词向量之前,还需对该告警日志进行文本分词。

[0111] 2) 计算该告警日志中每个词对应的特征权重；

[0112] 由于告警日志中有许多为了统一告警规范而存在的格式词,这些词在很多告警日志中都会出现,为了降低这些词对告警日志文本向量化特征表示的影响,有必要计算该告警日志中每个词对应的特征权重。而若一个词在该告警日志经常出现,而在其他告警日志很少出现,则说明该词对于该告警日志而言具有区分能力,有利于将该告警日志与其他告警日志进行区分。在一个具体的实施例,可以采用TFIDF方法(term frequency-inverse document frequency词频-逆文件频率)计算该告警日志中每个词对应的特征权重,具体地,采用TFIDF方法计算该告警日志中每个词对应的特征权重具体可以基于以下公式：

$$[0113] \quad \text{词频(TF)} = \frac{\text{该词在该告警日志中出现的次数}}{\text{该告警日志的词总数}}$$

$$[0114] \quad \text{逆文档频率(IDF)} = \log \left(\frac{\text{预设告警日志库的日志总数}}{\text{包含该词的告警日志数} + 1} \right)$$

[0115] TF-IDF值=词频(TF) × 逆文档频率(IDF)

[0116] 具体地,在逆文档频率(IDF)的计算公式中,对数函数的底数可以结合实际应用需求进行设定。上述的TF-IDF值可以表征该词对应的特征权重。

[0117] 3) 基于该告警日志中每个词对应的词向量和对应的特征权重进行加权求和,得到该告警日志对应的告警日志文本特征。

[0118] 通过基于预设词向量模型获取该告警日志中每个词对应的词向量,计算该告警日志中每个词对应的特征权重,基于该告警日志中每个词对应的词向量和对应的特征权重进行加权求和,得到该告警日志对应的告警日志文本特征,有利于降低无义词对告警日志文本特征的影响,确定出具有区分能力的词进行相应的权重设置,进而能够得到对于故障根因分析更加有利的告警日志文本特征,提升故障根因分析的准确性。

[0119] S209:基于根因关联概率分析模型,根据上述告警日志文本特征和上述隐含序列模式特征,对上述待分析组件集中组件进行故障根因关联概率分析,得到上述待分析组件集中组件间的故障根因关联概率。

[0120] 在本说明书实施例中,如图5所示,该根因关联概率分析模型可以包括相关性挖掘模块510、特征融合层520、前馈层530和分类层540。

[0121] 如图6所示,基于根因关联概率分析模型,根据上述告警日志文本特征和上述隐含序列模式特征,对上述待分析组件集中组件进行故障根因关联概率分析,得到上述待分析组件集中组件间的故障根因关联概率可以包括：

[0122] S601:基于上述的相关性挖掘模块对上述的告警日志文本特征进行相关性挖掘,得到告警日志相关性特征；

[0123] 由于指定时间范围内出现的告警日志往往具有很强的相关性,这种相关性对于故

障根因分析而言极为重要,因此有必要基于上述的相关性挖掘模块对上述的告警日志文本特征进行相关性挖掘。

[0124] 在本说明书实施例中,该相关性挖掘模块可以包括Transformer模型(基于自注意力机制的翻译模型),在实际应用中,该相关性挖掘模块可以作为该根因关联概率分析模型中的一部分,也可以作为一个独立的神经网络与根因关联概率分析模型级联。相较于CNN网络(Convolutional Neural Networks卷积神经网络)而言,Transformer模型更能够获取全局信息;相较于RNN网络(Recurrent Neural Network循环神经网络)而言,Transformer模型训练较快,效率高,且能够利用自注意力机制实现快速并行。在一个具体的实施例中,上述的Transformer模型可以包括但不限于多头自注意力模块、求和与归一化模块和前馈模块,其中,上述的多头自注意力模块可以由多个结构相同但权重矩阵不同的自注意力单元组成,从而使得每个自注意力单元能够关注到不同的特征,进而使得Transformer模型可以关注到更多的特征,避免模型只关注到部分特征的情况发生,有利于更全面地对上述告警日志文本特征进行相关性挖掘,得到更加准确的告警日志相关性特征,进而提升故障根因分析的准确性。

[0125] S603:基于上述的特征融合层对该告警日志相关性特征和上述的隐含序列模式特征进行特征融合,得到目标融合特征;

[0126] 在本说明书实施例中,利用该特征融合层能够对该告警日志相关性特征和上述的隐含序列模式特征进行深度特征抽取以实现特征融合,在一个实施例中,该特征融合层可以包括GRU层(Gate Recurrent Unit门控循环单元),相比于LSTM(Long-Short Term Memory长短期记忆网络)而言,GRU参数更少,且能够较好地处理序列信息;在另一个实施例中,该特征融合层也可以包括多个级联的前馈层,也能够有效地对该告警日志相关性特征和上述的隐含序列模式特征进行处理融合,本申请并不以此为限。

[0127] 请参照图7,当该相关性挖掘模块包括Transformer模型5101,该特征融合层包括GRU层5102时,该根因关联概率分析模型的结构如图7所示。

[0128] S605:基于上述的前馈层对该目标融合特征进行特征处理,得到处理后的目标融合特征;

[0129] 在本说明书实施例中,基于上述的前馈层对该目标融合特征进行特征处理可以包括但不限于对该目标融合特征进行特征抽取和权重配置。

[0130] S607:基于上述的分类层对该处理后的目标融合特征进行故障根因关联概率计算,得到该待分析组件集中组件间的故障根因关联概率。

[0131] 在本说明书实施例中,上述的分类层可以包括二分类层,可以将该待分析组件集中组件间是否存在故障根因关联作为概率问题,基于上述的分类层进行故障根因关联概率计算,输出该待分析组件集中组件间的故障根因关联概率,该故障根因关联概率可以表征该待分析组件集中组件间存在故障根因关联的概率。

[0132] 在本说明书实施例中,上述根因关联概率分析模型的损失函数可以包括但不限于交叉熵损失、铰链损失。

[0133] 通过基于根因关联概率分析模型,根据上述告警日志文本特征和上述隐含序列模式特征,对上述待分析组件集中组件进行故障根因关联概率分析,得到上述待分析组件集中组件间的故障根因关联概率,能够准确高效地确定几个组件间的故障根因关联概率,大

大降低对人工的依赖性,减少资源消耗。

[0134] S211:根据该待分析组件集中组件间的故障根因关联概率确定该待分析组件集中组件间的故障根因关联关系。

[0135] 具体地,上述的待分析组件集中组件间的故障根因关联关系可以包括:该待分析组件集中组件间存在故障关联,或,该待分析组件集中组件间不存在故障关联。

[0136] 请参照图8,在本说明书实施例中,根据该待分析组件集中组件间的故障根因关联概率确定该待分析组件集中组件间的故障根因关联关系可以包括:

[0137] S801:当该故障根因关联概率满足预设条件时,确定该待分析组件集中组件间的故障根因关联关系为存在故障根因关联。

[0138] 在一个具体的实施例中,上述的该故障根因关联概率满足预设条件可以包括该故障根因关联概率大于预设阈值,该预设阈值可以结合实际故障根因分析需求进行确定,例如,该预设阈值可以包括50%或80%。相应的,当该故障根因关联概率不满足预设条件时(例如该故障根因关联概率小于或等于预设阈值时),可以确定该待分析组件集中组件间的故障根因关联关系为不存在故障根因关联。

[0139] 通过根据该待分析组件集中组件间的故障根因关联概率确定该待分析组件集中组件间的故障根因关联关系,有利于能够帮助运维人员追溯故障的来源以及确定故障发生的相关因素,以使运维人员后续进行相应的维护,避免再次发生类似故障,减少故障带来的损失。

[0140] 在本说明书实施例中,还可以包括一种根因关联概率分析模型的训练方法,如下所示:

[0141] 1) 获取多个标注有组件间的故障根因关联概率的样本组件集对应的样本隐含序列模式特征,以及对应的样本告警日志文本特征;

[0142] 在本说明书实施例中,作为训练样本的样本组件集可以包括组件间存在故障根因关联的样本组件集(即组件间的故障根因关联概率较高),以及组件间不存在故障根因关联的样本组件集(即组件间的故障根因关联概率较低)。

[0143] 具体地,上述的获取多个标注有组件间的故障根因关联概率的样本组件集对应的样本告警日志文本特征可以包括:

[0144] 分别获取该样本组件集中的每个组件在第三预设时间范围内的样本告警日志,其中,该第三预设时间可以结合实际应用需求进行确定;分别对该样本告警日志进行文本向量化,得到对应的样本告警日志文本特征,具体过程与S207的过程类似,可以参见S207的相关描述,在此不再赘述。上述的获取多个标注有组件间的故障根因关联概率的样本组件集对应的样本隐含序列模式特征的具体过程与S201~S203的过程类似,可以参见S201~S203的相关描述,在此不再赘述。

[0145] 2) 基于上述的多个标注有组件间的故障根因关联概率的样本组件集对应的样本隐含序列模式特征,以及对应的样本告警日志文本特征对预设神经网络模型进行故障根因关联概率分析的训练,在故障根因关联概率分析的训练中调整该预设神经网络模型的模型参数至该预设神经网络模型满足预设收敛条件,得到上述故障根因关联概率分析模型。

[0146] 通过利用多个标注有组件间的故障根因关联概率的样本组件集对应的样本隐含序列模式特征,以及对应的样本告警日志文本特征对预设神经网络模型进行故障根因关联

概率分析的训练,有利于得到更加可靠的故障根因关联概率模型,进而提升故障根因分析的可靠性。

[0147] 由以上本说明书实施例提供的技术方案可见,本说明书实施例通过获取待分析组件集对应的多个待分析指标的原始时序信息,基于上述多个待分析指标的原始时序信息确定隐含序列模式特征,其中,基于上述多个待分析指标的原始时序信息确定隐含序列模式特征可以包括根据上述多个待分析指标的原始时序信息确定第二预设时间范围内的指标时序升降序列,根据上述指标时序升降序列进行序列模式挖掘,得到隐含序列模式,该隐含序列模式可以为上述多个待分析指标变化隐含的规律,可能是几个指标变化的关联关系或因果关系,后续可以结合各组件的告警日志进行故障根因分析,有利于提升故障根因分析的可靠性;由于各指标数据会随时间不断地更新变化,该隐含序列模式也会不断变化,某几个指标在过去一段时间没有关联,但之后可能会存在关联,因此可以根据需求调整上述的第二预设时间范围,来实时挖掘最新的隐含序列模式,灵活性强,提升故障根因分析的时效性;然后,对该隐含序列模式进行特征编码,得到隐含序列模式特征,获取该待分析组件集中的每个组件在第一预设时间范围内的告警日志,确定与上述每个组件在第一预设时间范围内的告警日志对应的告警日志文本特征;然后,基于根因关联概率分析模型,根据上述告警日志文本特征和上述隐含序列模式特征,对该待分析组件集中组件进行故障根因关联概率分析,得到该待分析组件集中组件间的故障根因关联概率,能够准确高效地确定几个组件间的故障根因关联概率,大大降低对人工的依赖性,减少资源消耗;最后,根据待分析组件集中组件间的故障根因关联概率确定待分析组件集中组件间的故障根因关联关系,有利于能够帮助运维人员追溯故障的来源以及确定故障发生的相关因素,以使运维人员后续进行相应的维护,避免再次发生类似故障,减少故障带来的损失。

[0148] 本申请实施例还提供了一种故障根因分析装置,如图9所示,所述装置可以包括:

[0149] 原始时序信息获取模块910,用于获取待分析组件集对应的多个待分析指标的原始时序信息,所述多个待分析指标包括所述待分析组件集中的每个组件对应的待分析指标;

[0150] 隐含序列模式特征确定模块920,用于基于所述多个待分析指标的原始时序信息确定隐含序列模式特征;

[0151] 告警日志获取模块930,用于获取所述待分析组件集中的每个组件在第一预设时间范围内的告警日志;

[0152] 文本特征确定模块940,用于确定与上述每个组件在第一预设时间范围内的告警日志对应的告警日志文本特征;

[0153] 故障根因关联概率分析模块950,用于基于根因关联概率分析模型,根据所述告警日志文本特征和所述隐含序列模式特征,对所述待分析组件集中组件进行故障根因关联概率分析,得到所述待分析组件集中组件间的故障根因关联概率;

[0154] 故障根因关联关系确定模块960,用于根据所述待分析组件集中组件间的故障根因关联概率确定所述待分析组件集中组件间的故障根因关联关系。

[0155] 在一些实施例中,上述的根因关联概率分析模型可以包括:

[0156] 相关性挖掘模块、特征融合层、前馈层和分类层。

[0157] 当上述的根因关联概率分析模型包括相关性挖掘模块、特征融合层、前馈层和分

类层时,上述的故障根因关联概率分析模块950可以包括:

[0158] 相关性挖掘单元,用于基于所述相关性挖掘模块对所述告警日志文本特征进行相关性挖掘,得到告警日志相关性特征;

[0159] 特征融合单元,用于基于所述特征融合层对所述告警日志相关性特征和所述隐含序列模式特征进行特征融合,得到目标融合特征;

[0160] 特征处理单元,用于基于所述前馈层对所述目标融合特征进行特征处理,得到处理后的目标融合特征;

[0161] 故障根因关联概率确定单元,用于基于所述分类层对所述处理后的目标融合特征进行故障根因关联概率计算,得到所述待分析组件集中组件间的故障根因关联概率。

[0162] 在一些实施例中,上述的隐含序列模式特征确定模块920可以包括:

[0163] 指标时序升降序列确定单元,用于根据所述多个待分析指标的原始时序信息确定第二预设时间范围内的指标时序升降序列;

[0164] 序列模式挖掘单元,用于根据所述指标时序升降序列进行序列模式挖掘,得到隐含序列模式;

[0165] 特征编码单元,用于对所述隐含序列模式进行特征编码,得到上述隐含序列模式特征。

[0166] 在一些实施例中,上述的指标时序升降序列确定单元可以包括:

[0167] 时序升降信息确定单元,用于根据所述多个待分析指标的原始时序信息确定所述多个待分析指标的时序升降信息;

[0168] 时序升降序列构造单元,用于根据所述多个待分析指标的时序升降信息构造所述第二预设时间范围内的指标时序升降序列。

[0169] 在一些实施例中,所述装置还可以包括:

[0170] 样本数据获取单元,用于获取多个标注有组件间的故障根因关联概率的样本组件集对应的样本隐含序列模式特征,以及对应的样本告警日志文本特征;

[0171] 模型训练单元,用于基于所述多个标注有组件间的故障根因关联概率的样本组件集对应的样本隐含序列模式特征,以及对应的样本告警日志文本特征对预设神经网络模型进行故障根因关联概率分析的训练,在故障根因关联概率分析的训练中调整所述预设神经网络模型的模型参数至所述预设神经网络模型满足预设收敛条件,得到所述故障根因关联概率分析模型。

[0172] 在一些实施例中,所述样本数据获取单元可以包括:

[0173] 样本告警日志获取单元,用于分别获取所述样本组件集中的每个组件在第三预设时间范围内的样本告警日志;

[0174] 文本向量化单元,用于分别对所述样本告警日志进行文本向量化,得到对应的样本告警日志文本特征。

[0175] 在一些实施例中,所述故障根因关联关系确定模块960可以包括:

[0176] 故障根因关联确定单元,用于当所述故障根因关联概率满足预设条件时,确定所述待分析组件集中组件间的故障根因关联关系为存在故障根因关联。

[0177] 所述的装置实施例中的装置与方法实施例基于同样地申请构思。

[0178] 本申请实施例提供了一种计算机设备,该计算机设备包括处理器和存储器,该存

存储器中存储有至少一条指令或至少一段程序,该至少一条指令或该至少一段程序由该处理器加载并执行以实现如上述方法实施例所提供的故障根因分析方法。

[0179] 存储器可用于存储软件程序以及模块,处理器通过运行存储在存储器的软件程序以及模块,从而执行各种功能应用以及数据处理。存储器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、功能所需的应用程序等;存储数据区可存储根据所述设备的使用所创建的数据等。此外,存储器可以包括高速随机存取存储器,还可以包括非易失性存储器,例如至少一个磁盘存储器件、闪存器件、或其他易失性固态存储器件。相应地,存储器还可以包括存储器控制器,以提供处理器对存储器的访问。

[0180] 本申请实施例所提供的方法实施例可以在移动终端、计算机终端、服务器或者类似的运算装置中执行,即上述计算机设备可以包括移动终端、计算机终端、服务器或者类似的运算装置。其中,上述的服务器可以是独立的物理服务器,也可以是多个物理服务器构成的服务器集群或者分布式系统,还可以是提供云服务、云数据库、云计算、云函数、云存储、网络服务、云通信、中间件服务、域名服务、安全服务、CDN、以及大数据和人工智能平台等基础云计算服务的云服务器。终端可以是智能手机、平板电脑、笔记本电脑、台式计算机、智能音箱、智能手表等,但并不局限于此。以运行在服务器上为例,图10是本申请实施例提供的一种用于实现上述故障根因分析方法的服务器的硬件结构框图。如图10所示,该服务器1000可因配置或性能不同而产生比较大的差异,可以包括一个或一个以上中央处理器(Central Processing Units,CPU)1010(处理器1010可以包括但不限于微处理器MCU或可编程逻辑器件FPGA等的处理装置)、用于存储数据的存储器1030,一个或一个以上存储应用程序1023或数据1022的存储介质1020(例如一个或一个以上海量存储设备)。其中,存储器1030和存储介质1020可以是短暂存储或持久存储。存储在存储介质1020的程序可以包括一个或一个以上模块,每个模块可以包括对服务器中的一系列指令操作。更进一步地,中央处理器1010可以设置为与存储介质1020通信,在服务器1000上执行存储介质1020中的一系列指令操作。服务器1000还可以包括一个或一个以上电源1060,一个或一个以上有线或无线网络接口1050,一个或一个以上输入输出接口1040,和/或,一个或一个以上操作系统1021,例如Windows Server™,Mac OS X™,Unix™,Linux™,FreeBSD™等等。

[0181] 处理器1010可以是一种集成电路芯片,具有信号的处理能力,例如通用处理器、数字信号处理器(DSP,Digital Signal Processor),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任何常规的处理器等。

[0182] 输入输出接口1040可以用于经由一个网络接收或者发送数据。上述的网络具体实例可包括服务器1000的通信供应商提供的无线网络。在一个实例中,输入输出接口1040包括一个网络适配器(Network Interface Controller,NIC),其可通过基站与其他网络设备相连从而可与互联网进行通讯。在一个实例中,输入输出接口1040可以为射频(Radio Frequency,RF)模块,其用于通过无线方式与互联网进行通讯。

[0183] 操作系统1021可以包括用于处理各种基本系统服务和执行硬件相关任务的系统程序,例如框架层、核心库层、驱动层等,用于实现各种基础业务以及处理基于硬件的任务。

[0184] 本领域普通技术人员可以理解,图10所示的结构仅为示意,其并不对上述电子装置的结构造成限定。例如,服务器1000还可包括比图10中所示更多或者更少的组件,或者具

有与图10所示不同的配置。

[0185] 本申请的实施例还提供了一种计算机可读存储介质,所述存储介质可设置于服务器之中以保存用于实现方法实施例中一种故障根因分析方法相关的至少一条指令或至少一段程序,该至少一条指令或该至少一段程序由该处理器加载并执行以实现上述方法实施例提供的故障根因分析方法。

[0186] 可选地,在本实施例中,上述存储介质可以位于计算机网络的多个网络服务器中的至少一个网络服务器。可选地,在本实施例中,上述存储介质可以包括但不限于:U盘、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、移动硬盘、磁碟或者光盘等各种可以存储程序代码的介质。

[0187] 由上述本申请提供的故障根因分析方法、装置、设备或存储介质的实施例可见,本申请中通过获取待分析组件集对应的多个待分析指标的原始时序信息,基于上述多个待分析指标的原始时序信息确定隐含序列模式特征,其中,基于上述多个待分析指标的原始时序信息确定隐含序列模式特征可以包括根据上述多个待分析指标的原始时序信息确定第二预设时间范围内的指标时序升降序列,根据上述指标时序升降序列进行序列模式挖掘,得到隐含序列模式,该隐含序列模式可以为上述多个待分析指标变化隐含的规律,可能是几个指标变化的关联关系或因果关系,后续可以结合各组件的告警日志进行故障根因分析,有利于提升故障根因分析的可靠性;由于各指标数据会随时间不断地更新变化,该隐含序列模式也会不断变化,某几个指标在过去一段时间没有关联,但之后可能会存在关联,因此可以根据需求调整上述的第二预设时间范围,来实时挖掘最新的隐含序列模式,灵活性强,提升故障根因分析的时效性;然后,对该隐含序列模式进行特征编码,得到隐含序列模式特征,获取该待分析组件集中的每个组件在第一预设时间范围内的告警日志,确定与上述每个组件在第一预设时间范围内的告警日志对应的告警日志文本特征;然后,基于根因关联概率分析模型,根据所述告警日志文本特征和所述隐含序列模式特征,对所述待分析组件集中组件进行故障根因关联概率分析,得到所述待分析组件集中组件间的故障根因关联概率,能够准确高效地确定几个组件间的故障根因关联概率,大大降低对人工的依赖性,减少资源消耗;最后,根据待分析组件集中组件间的故障根因关联概率确定待分析组件集中组件间的故障根因关联关系,有利于能够帮助运维人员追溯故障的来源以及确定故障发生的相关因素,以使运维人员后续进行相应的维护,避免再次发生类似故障,减少故障带来的损失。

[0188] 需要说明的是:上述本申请实施例先后顺序仅仅为了描述,不代表实施例的优劣。且上述对本说明书特定实施例进行了描述。其它实施例在所附权利要求书的范围内。在一些情况下,在权利要求书中记载的动作或步骤可以按照不同于实施例中的顺序来执行并且仍然可以实现期望的结果。另外,在附图中描绘的过程不一定要示出的特定顺序或者连续顺序才能实现期望的结果。在某些实施方式中,多任务处理和并行处理也是可以的或者可能是有利的。

[0189] 本说明书中的各个实施例均采用递进的方式描述,各个实施例之间相同相似的部分互相参见即可,每个实施例重点说明的都是与其他实施例的不同之处。尤其,对于装置、设备和存储介质实施例而言,由于其基本相似于方法实施例,所以描述的比较简单,相关之处参见方法实施例的部分说明即可。

[0190] 本领域普通技术人员可以理解实现上述实施例的全部或部分步骤可以通过硬件来完成,也可以通过程序来指令相关的硬件完成,所述的程序可以存储于一种计算机可读存储介质中,上述提到的存储介质可以是只读存储器,磁盘或光盘等。

[0191] 以上所述仅为本申请的较佳实施例,并不用以限制本申请,凡在本申请的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本申请的保护范围之内。

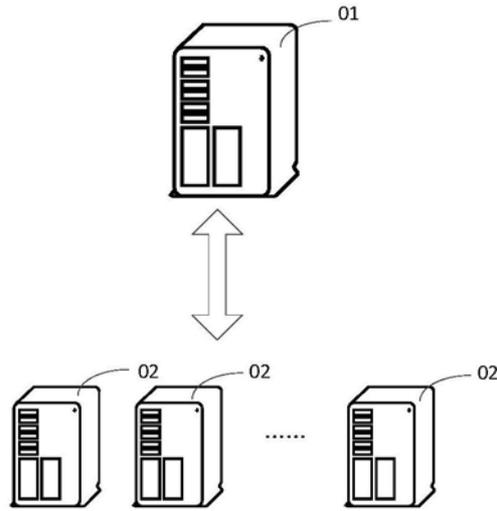


图1

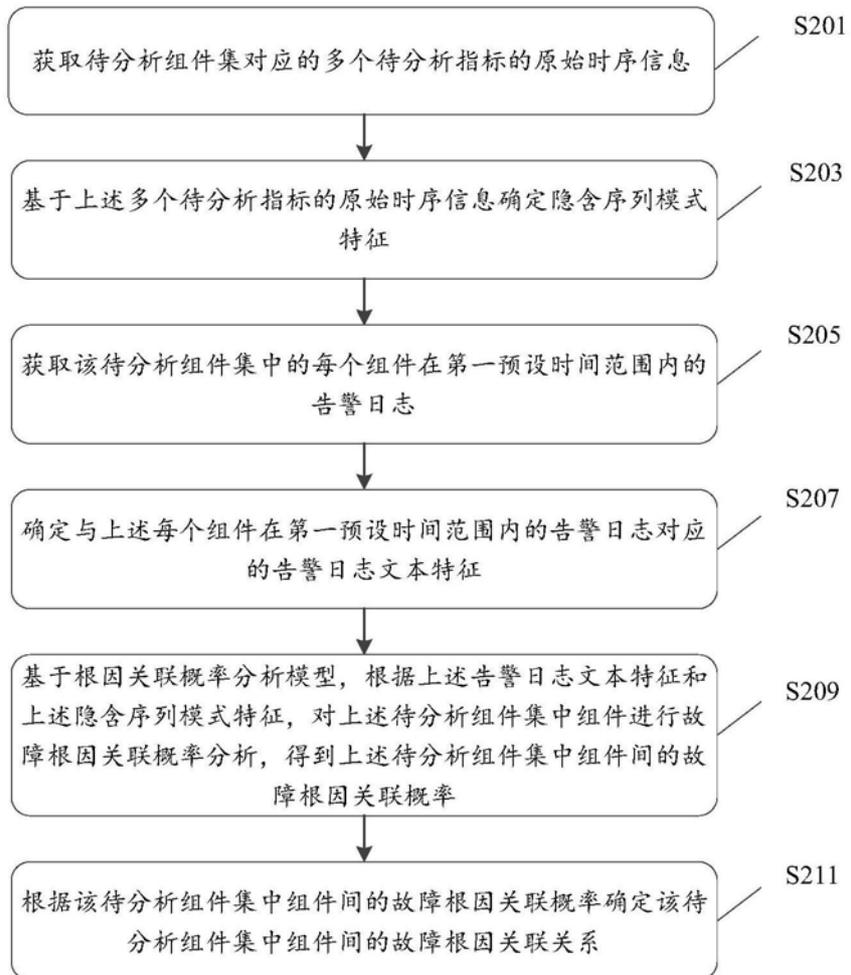


图2

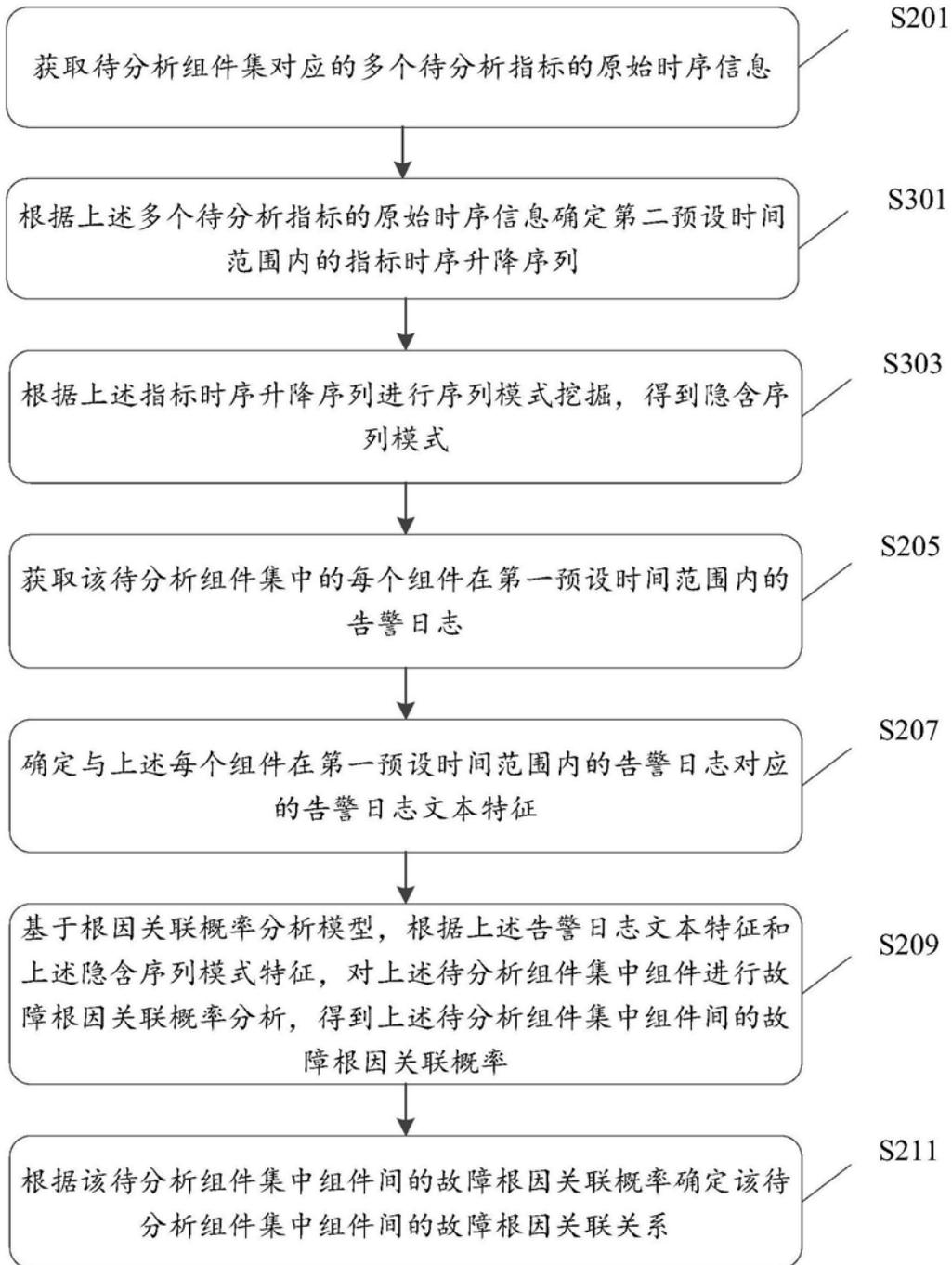


图3

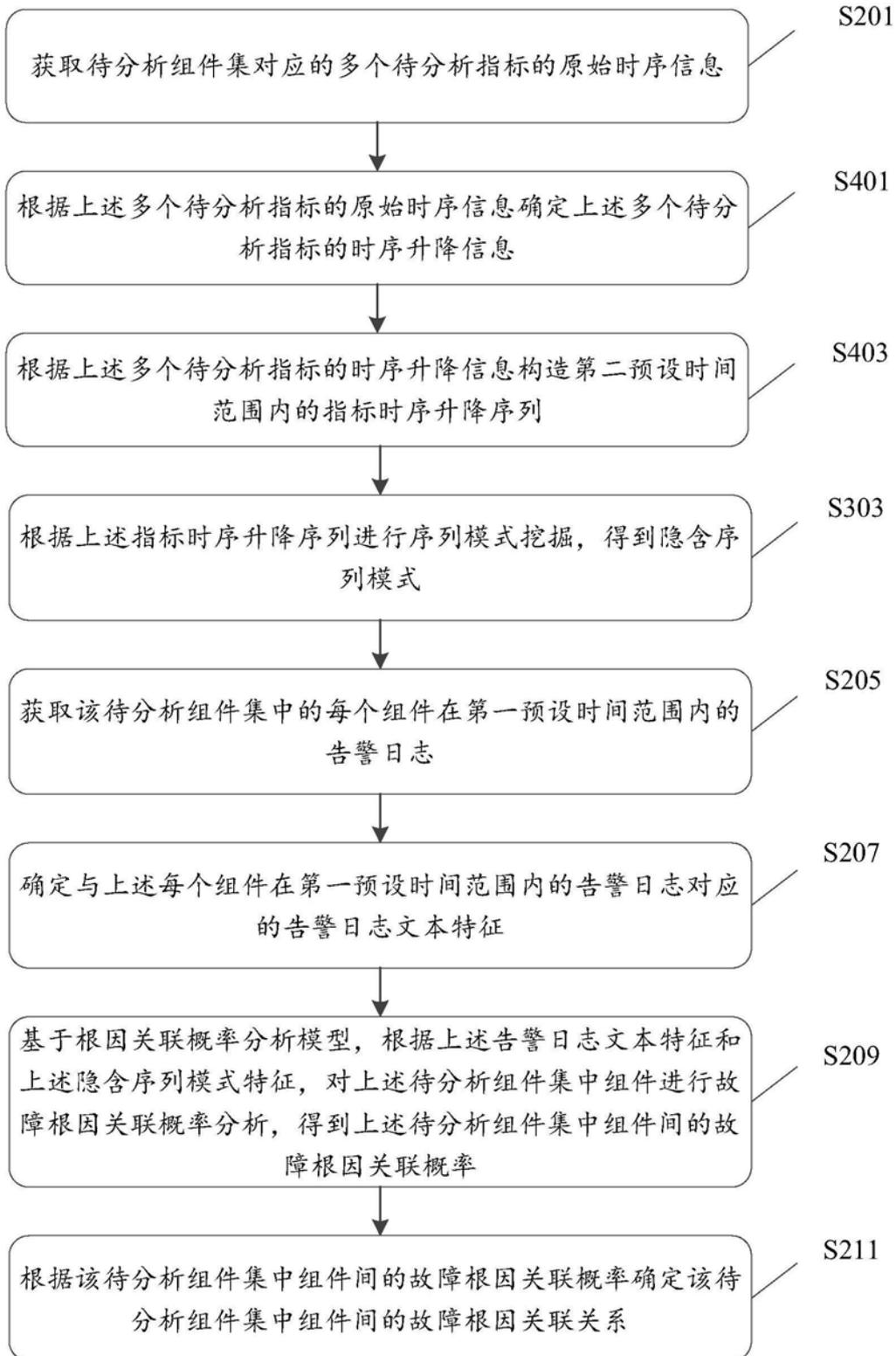


图4

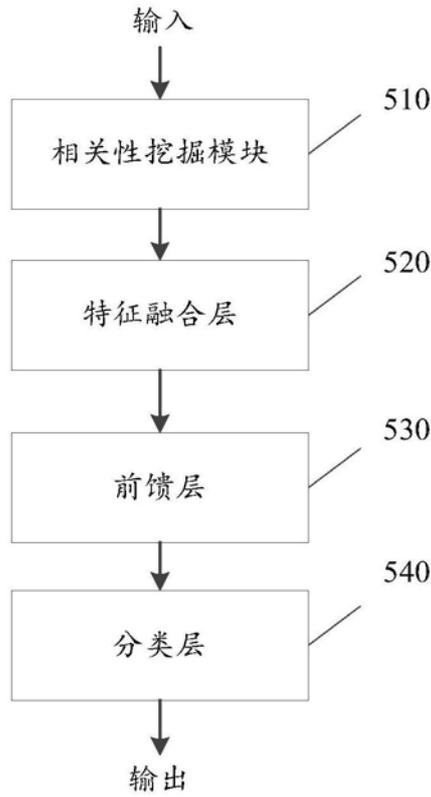


图5

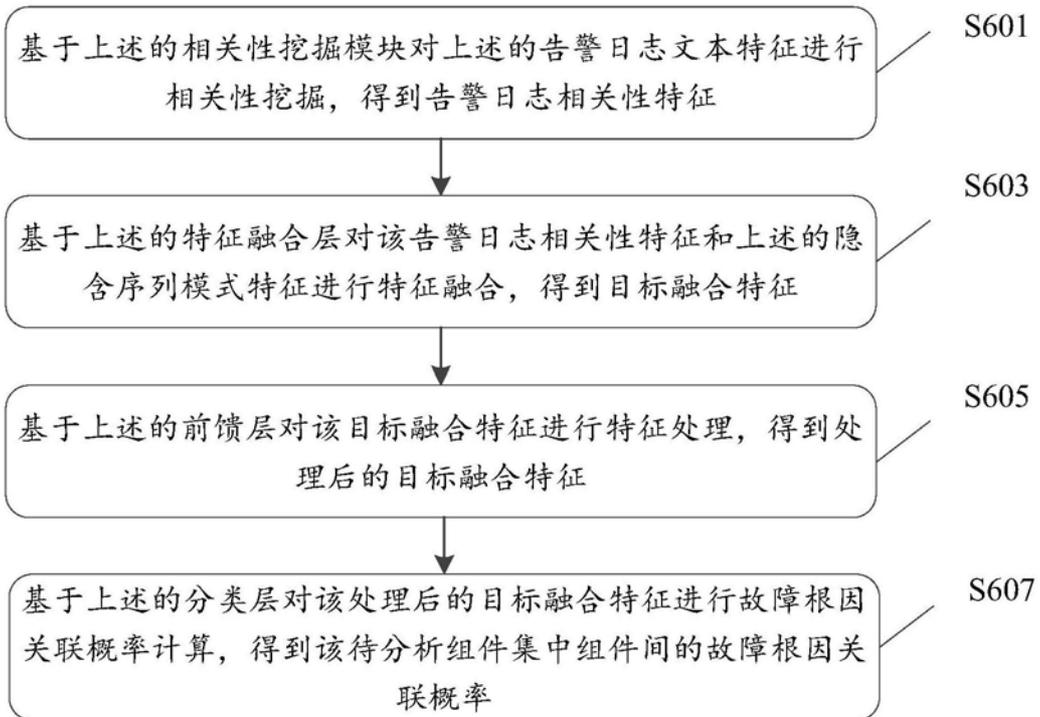


图6

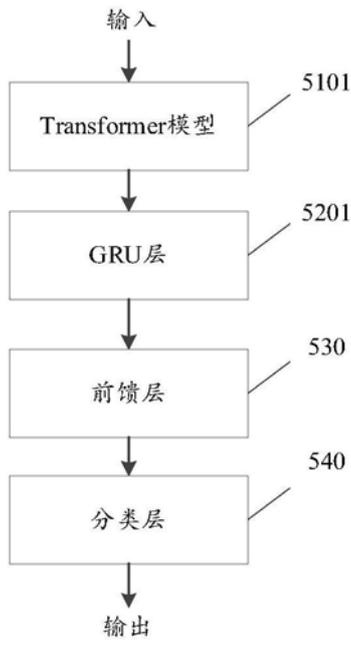


图7

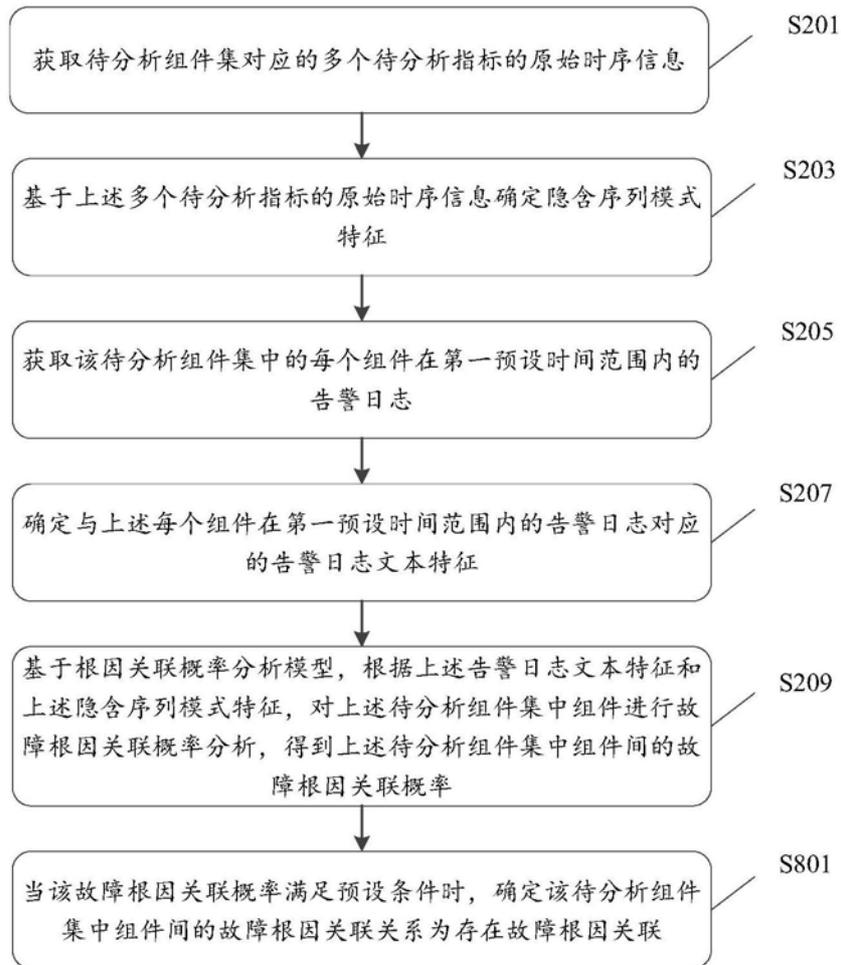


图8

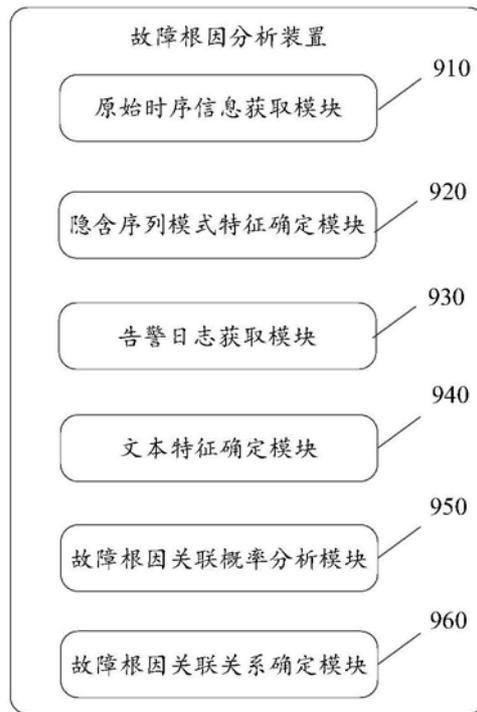


图9

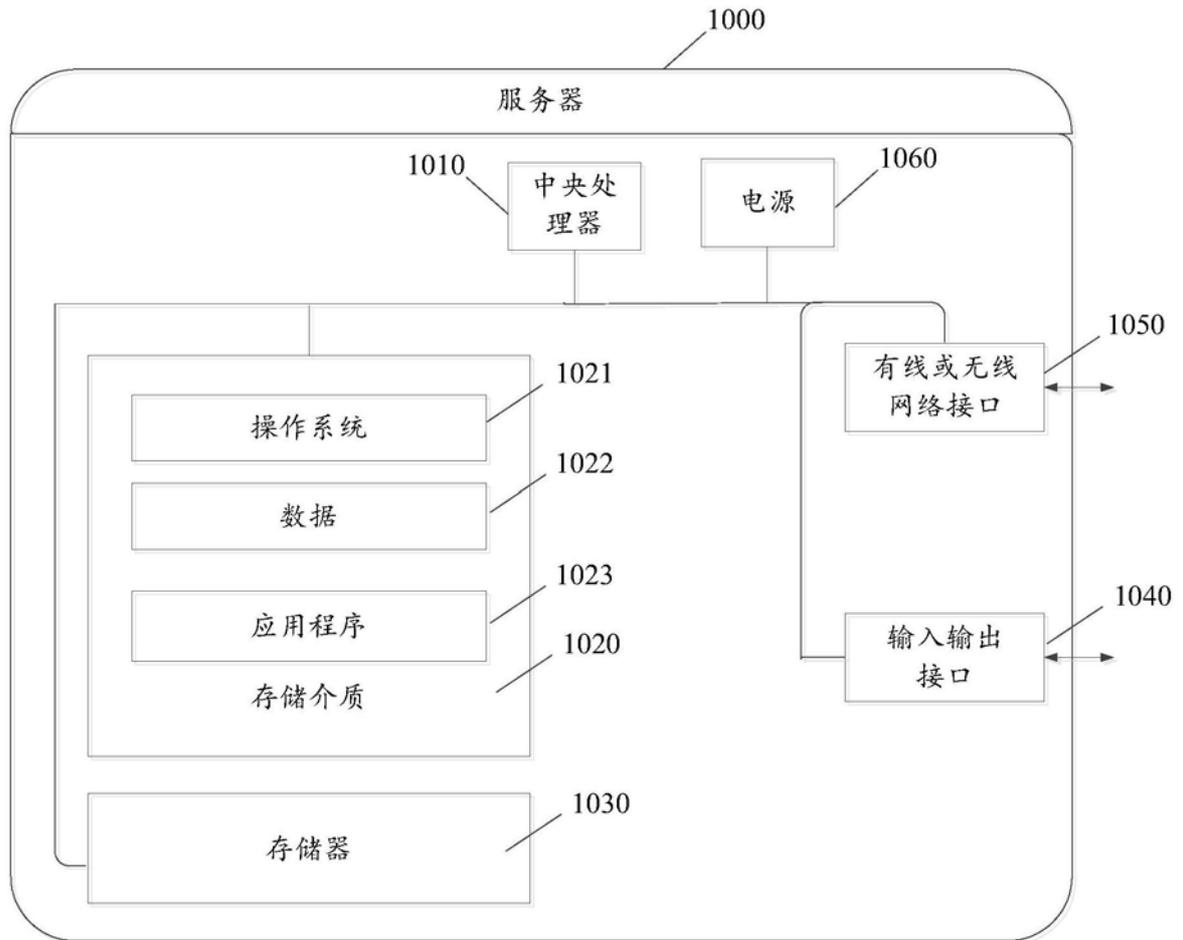


图10