



(12)发明专利申请

(10)申请公布号 CN 111160329 A

(43)申请公布日 2020.05.15

(21)申请号 201911381777.6

(22)申请日 2019.12.27

(71)申请人 深圳前海微众银行股份有限公司
地址 518027 广东省深圳市前海深港合作区前湾一路1号A栋201室

(72)发明人 杜林 尤嘉 李诚 林城 刘超 杨利国

(74)专利代理机构 北京同达信恒知识产权代理有限公司 11291

代理人 彭燕

(51)Int.Cl.

G06K 9/03(2006.01)

G06Q 40/00(2012.01)

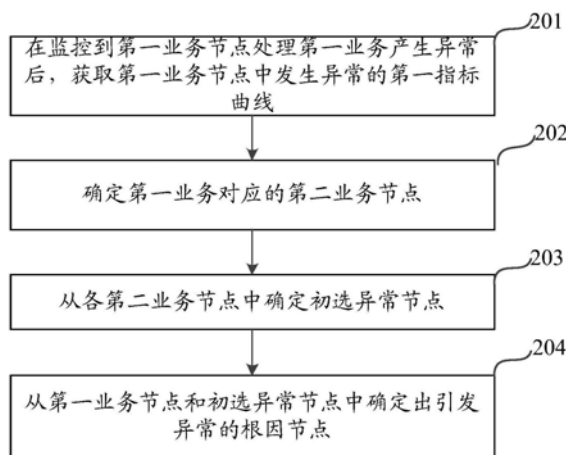
权利要求书2页 说明书11页 附图3页

(54)发明名称

一种根因分析的方法及装置

(57)摘要

本发明实施例公开了一种根因分析的方法及装置,方法包括:在监控到第一业务节点处理第一业务产生异常后,获取第一业务节点中发生异常的第一指标曲线,确定第一业务对应的第二业务节点,从各第二业务节点中确定初选异常节点,其中,初选异常节点中具有与第一指标曲线匹配的第二指标曲线,从第一业务节点和初选异常节点中确定出引发异常的根因节点。本发明通过将具有与第一指标曲线匹配的第二指标曲线的节点作为初选异常节点也即根据曲线走势判断相似性,高效进行根因分析的同时提高了根因分析结果的准确性。



1. 一种根因分析的方法,其特征在于,包括:

在监控到第一业务节点处理第一业务产生异常后,获取所述第一业务节点中发生异常的第一指标曲线;

确定所述第一业务对应的第二业务节点,所述第二业务节点为处理所述第一业务的各业务节点中除所述第一业务节点之外的任一业务节点;

从各第二业务节点中确定初选异常节点,其中,所述初选异常节点中具有与所述第一指标曲线匹配的第二指标曲线;

从所述第一业务节点和所述初选异常节点中确定出引发异常的根因节点。

2. 根据权利要求1所述的方法,其特征在于,所述从各第二业务节点中确定初选异常节点,包括:

针对每个第二业务节点,确定所述第二业务节点各指标曲线中与所述第一业务关联的指标曲线;若与所述第一业务关联的指标曲线中存在与所述第一指标曲线匹配的第二指标曲线,则确定所述第二业务节点为初选异常节点。

3. 根据权利要求1所述的方法,其特征在于,在所述在监控到第一业务节点处理第一业务产生异常之前,还包括:

通过第三业务的业务处理流水,确定处理所述第三业务的各第三业务节点,并确定每个第三业务节点各指标曲线中与所述第三业务关联的指标曲线;所述第三业务为各业务中的任一种;

生成所述第三业务的业务标识、各第三业务节点、各第三业务节点中与所述第三业务关联的指标曲线的对照关系。

4. 根据权利要求1所述的方法,其特征在于,通过如下方式确定与所述第一指标曲线是否匹配的第二指标曲线,包括:

确定所述第一指标曲线中的第一异常开始时刻、第一异常峰值时刻及异常时段中各时刻的指标波动值的第一排序;

确定所述第二指标曲线中的第二异常开始时刻、第二异常峰值时刻及异常时段中各时刻的指标波动值的第二排序;

根据所述第一异常开始时刻与所述第二异常开始时刻的第一匹配度、所述第一异常峰值时刻与所述第二异常峰值时刻的第二匹配度及所述第一排序与所述第二排序的第三匹配度,确定所述第一指标曲线与所述第二指标曲线之间的匹配度是否大于设定阈值。

5. 根据权利要求1所述的方法,其特征在于,所述从所述第一业务节点和所述初选异常节点中确定出引发异常的根因节点,包括:

针对发生异常的同—指标,确定所述第一业务节点在所述指标的第一变动量,及所述初选异常节点在所述指标的第二变动量;

若所述第一变动量与所述第二变动量之间的关系不满足设定条件,则排除所述初选异常节点;

从排除操作后的初选异常节点和所述第一业务节点中确定出引发异常的根因节点。

6. 根据权利要求5所述的方法,其特征在于,从排除操作后的初选异常节点和所述第一业务节点中确定出引发异常的根因节点,包括:

针对所述第一指标曲线对应的指标,从所述初选异常节点和所述第一业务节点中确定

出在处理所述第一业务时在所述指标上的影响方,将影响方确定为引发异常的根因节点。

7. 根据权利要求1至6任一项所述的方法,其特征在于,所述方法还包括:

确定所述第一指标曲线对应的指标所关联的第一属性及所述第二指标曲线对应的指标所关联的第二属性;

展示所述第一属性和所述第二属性中的相同属性。

8. 一种根因分析的装置,其特征在于,所述装置包括:

获取模块,用于在监控到第一业务节点处理第一业务产生异常后,获取所述第一业务节点中发生异常的第一指标曲线;

处理模块,用于确定所述第一业务对应的第二业务节点,所述第二业务节点为处理所述第一业务的各业务节点中除所述第一业务节点之外的任一业务节点;

所述处理模块,还用于从各第二业务节点中确定初选异常节点,其中,所述初选异常节点中具有与所述第一指标曲线匹配的第二指标曲线;

所述处理模块,还用于从所述第一业务节点和所述初选异常节点中确定出引发异常的根因节点。

9. 一种计算设备,其特征在于,包括:

存储器,用于存储程序指令;

处理器,用于调用所述存储器中存储的程序指令,按照获得的程序执行权利要求1至7任一项所述的方法。

10. 一种计算机可读非易失性存储介质,其特征在于,包括计算机可读指令,当计算机读取并执行所述计算机可读指令时,使得计算机执行如权利要求1至7任一项所述的方法。

一种根因分析的方法及装置

技术领域

[0001] 本发明涉及金融科技 (Fintech) 计算机领域, 尤其涉及一种根因分析的方法及装置。

背景技术

[0002] 根源问题分析 (Root Cause Analysis, RCA) 是指在系统发生异常的时候会出现很多相关的异常信息, 从这些大量相关的信息中找到问题的根源。RCA的分析, 是通过系统体现出来的一些数据和数据之间的关系, 达到找寻问题根源的目的。比如: 某主机磁盘异常, 某系统的某一命令执行时间过长, 某方法出现了空指针报错等。

[0003] RCA一直没有通用的解决方法。当系统达到一定的规模时, 会出现子系统多、关系多、指标多、告警多的情况, 这个时候来进行RCA分析会非常困难。

[0004] 目前常用的RCA方案为告警分析: 当出现业务异常的时候, 完善的监控系统一般都会有大量的监控告警。这些告警有一些是基于指标的阈值, 有一些是基于错误关键字日志, 有一些是基于系统的模块调用等。这些告警一般均是以自身采集的数据来作为监控点, 当发生业务异常时, 有可能只是受到其它系统异常的影响, 不一定是问题的根源。也就是说, 业务异常时, 可能会有大量的告警, 可能其中一些是根源, 另一些可能和该异常有关, 但不是问题的根源。基于告警的RCA分析, 就是要从众多的告警中, 去除高频告警和无用告警, 并梳理出依赖关系, 最终判断问题的根源。

[0005] 上述方案所提出的告警分析更多的是从微观的角度来看待问题的发生。比如: 一笔交易中断了, 某一个接口报错了, 输出了失败日志等。这种从微观的角度来分析问题, 需要进行大量的分析, 从而使得在异常出现时发现根因的效率很低。

发明内容

[0006] 本发明实施例提供一种根因分析的方法及装置, 用以解决如何高效准确的进行根因分析的问题。

[0007] 第一方面, 本发明实施例提供一种根因分析的方法, 包括:

[0008] 在监控到第一业务节点处理第一业务产生异常后, 获取所述第一业务节点中发生异常的第一指标曲线;

[0009] 确定所述第一业务对应的第二业务节点, 所述第二业务节点为处理所述第一业务的各业务节点中除所述第一业务节点之外的任一业务节点;

[0010] 从各第二业务节点中确定初选异常节点, 其中, 所述初选异常节点中具有与所述第一指标曲线匹配的第二指标曲线;

[0011] 从所述第一业务节点和所述初选异常节点中确定出引发异常的根因节点。

[0012] 上述方案, 在监控到业务节点发生异常后, 通过将发生异常的第一指标曲线和处理第一业务的其他业务节点的指标曲线进行匹配, 将具有与第一指标曲线匹配的第二指标曲线的业务节点作为初选异常节点也即根据曲线走势判断相似性, 再从初选异常节点和第

一业务节点中确定根因节点,可以在高效进行根因分析的同时提高了根因分析结果的准确性。

[0013] 可选的,所述从各第二业务节点中确定初选异常节点,包括:

[0014] 针对每个第二业务节点,确定所述第二业务节点各指标曲线中与所述第一业务关联的指标曲线;若与所述第一业务关联的指标曲线中存在与所述第一指标曲线匹配的第二指标曲线,则确定所述第二业务节点为初选异常节点。

[0015] 上述方案,从第一业务对应的第二业务节点各指标曲线中确定出与所述第一业务关联的指标曲线,将与所述第一业务关联的指标曲线进行匹配;由于实际业务中,一个业务节点会涉及多种业务,而且一个业务节点上会基于各种需求设置很多的指标曲线;本方案中通过缩小业务节点的范围,缩小指标曲线的范围,可以提高了根因分析的效率;同时,由于排除了不相关的曲线,从而避免了某些曲线由于波动导致了与第一指标曲线的匹配度提升,从而也提高了根因分析的准确度。

[0016] 可选的,在所述在监控到第一业务节点处理第一业务产生异常之前,还包括:

[0017] 通过第三业务的业务处理流水,确定处理所述第三业务的各第三业务节点,并确定每个第三业务节点各指标曲线中与所述第三业务关联的指标曲线;所述第三业务为各业务中的任一种;

[0018] 生成所述第三业务的业务标识、各第三业务节点、各第三业务节点中与所述第三业务关联的指标曲线的对照关系。

[0019] 上述方案,避免了由于相似判断考察的范围过大,本方案中梳理出每个业务在处理时经过的业务节点及该业务节点上涉及的指标曲线,从而在做相似判定的时候,则只在对照关系中查找该业务涉及的业务节点和指标曲线就行。这样判断相似的集合大大缩小,准确性得到提高。

[0020] 可选的,通过如下方式确定与所述第一指标曲线是否匹配的第二指标曲线,包括:

[0021] 确定所述第一指标曲线中的第一异常开始时刻、第一异常峰值时刻及异常时段中各时刻的指标波动值的第一排序;

[0022] 确定所述第二指标曲线中的第二异常开始时刻、第二异常峰值时刻及异常时段中各时刻的指标波动值的第二排序;

[0023] 根据所述第一异常开始时刻与所述第二异常开始时刻的第一匹配度、所述第一异常峰值时刻与所述第二异常峰值时刻的第二匹配度及所述第一排序与所述第二排序的第三匹配度,确定所述第一指标曲线与所述第二指标曲线之间的匹配度是否大于设定阈值。

[0024] 上述方案,通过从异常开始时刻、异常峰值时刻及指标波动值的排序三个维度进行匹配度的观察,可以得到某一条曲线和发生异常的指标曲线的相似程度。当高于一定的阈值时,将其确定为初选异常节点,通过匹配度的计算,提高了根因分析结果的准确性。

[0025] 可选的,所述从所述第一业务节点和所述初选异常节点中确定出引发异常的根因节点,包括:

[0026] 针对发生异常的第一指标,确定所述第一业务节点在所述第一指标的第一变动量,及所述初选异常节点在所述第一指标的第二变动量;

[0027] 若所述第一变动量与所述第二变动量之间的关系不满足设定条件,则排除所述初选异常节点;

[0028] 从排除操作后的初选异常节点和所述第一业务节点中确定出引发异常的根因节点。

[0029] 上述方法,通过第一变动量与第二变动量之间的关系可以排除掉一些原本被判定相似,但不合理的曲线,提高判定的准确性。

[0030] 可选的,从排除操作后的初选异常节点和所述第一业务节点中确定出引发异常的根因节点,包括:

[0031] 针对所述第一指标曲线对应的指标,从所述初选异常节点和所述第一业务节点中确定出在处理所述第一业务时在所述指标上的影响方,将影响方确定为引发异常的根因节点。

[0032] 上述方案,根据初选异常节点与发生异常节点之间的依赖关系确定最终的根因节点,减少了误判的可能,提高了根因分析的准确性。

[0033] 可选的,所述方法还包括:

[0034] 确定所述第一指标曲线对应的指标所关联的第一属性及所述第二指标曲线对应的指标所关联的第二属性;

[0035] 展示所述第一属性和所述第二属性中的相同属性。

[0036] 上述方案,通过找到所有相似异动的指标及指标所关联的属性并进行统计和输出,可以将异常圈定在一个更小的范围或分类中,提高了根因分析结果的可读性。

[0037] 第二方面,本发明实施例提供一种根因分析的装置,包括:

[0038] 获取模块,用于在监控到第一业务节点处理第一业务产生异常后,获取所述第一业务节点中发生异常的第一指标曲线;

[0039] 处理模块,用于确定所述第一业务对应的第二业务节点,所述第二业务节点为处理所述第一业务的各业务节点中除所述第一业务节点之外的任一业务节点;

[0040] 所述处理模块,还用于从各第二业务节点中确定初选异常节点,其中,所述初选异常节点中具有与所述第一指标曲线匹配的第二指标曲线;

[0041] 所述处理模块,还用于从所述第一业务节点和所述初选异常节点中确定出引发异常的根因节点。

[0042] 可选的,所述处理模块具体用于:

[0043] 针对每个第二业务节点,确定所述第二业务节点各指标曲线中与所述第一业务关联的指标曲线;若与所述第一业务关联的指标曲线中存在与所述第一指标曲线匹配的第二指标曲线,则确定所述第二业务节点为初选异常节点。

[0044] 可选的,所述处理模块还用于:

[0045] 在所述在监控到第一业务节点处理第一业务产生异常之前,通过第三业务的业务处理流水,确定处理所述第三业务的各第三业务节点,并确定每个第三业务节点各指标曲线中与所述第三业务关联的指标曲线;所述第三业务为各业务中的任一种;

[0046] 生成所述第三业务的业务标识、各第三业务节点、各第三业务节点中与所述第三业务关联的指标曲线的对照关系。

[0047] 可选的,所述处理模块具体用于:

[0048] 确定所述第一指标曲线中的第一异常开始时刻、第一异常峰值时刻及异常时段中各时刻的指标波动值的第一排序;

- [0049] 确定所述第二指标曲线中的第二异常开始时刻、第二异常峰值时刻及异常时段中各时刻的指标波动值的第二排序；
- [0050] 根据所述第一异常开始时刻与所述第二异常开始时刻的第一匹配度、所述第一异常峰值时刻与所述第二异常峰值时刻的第二匹配度及所述第一排序与所述第二排序的第三匹配度，确定所述第一指标曲线与所述第二指标曲线之间的匹配度是否大于设定阈值。
- [0051] 可选的，所述处理模块具体用于：
- [0052] 针对发生异常的第一指标，确定所述第一业务节点在所述第一指标的第一变动量，及所述初选异常节点在所述第二指标的第二变动量；
- [0053] 若所述第一变动量与所述第二变动量之间的关系不满足设定条件，则排除所述初选异常节点；
- [0054] 从排除操作后的初选异常节点和所述第一业务节点中确定出引发异常的根因节点。
- [0055] 可选的，所述处理模块具体用于：
- [0056] 针对所述第一指标曲线对应的指标，从所述初选异常节点和所述第一业务节点中确定出在处理所述第一业务时在所述指标上的影响方，将影响方确定为引发异常的根因节点。
- [0057] 可选的，所述处理模块还用于：
- [0058] 确定所述第一指标曲线对应的指标所关联的第一属性及所述第二指标曲线对应的指标所关联的第二属性；
- [0059] 展示所述第一属性和所述第二属性中的相同属性。
- [0060] 第三方面，本发明实施例提供一种计算机被控设备，包括：
- [0061] 存储器，用于存储程序指令；
- [0062] 处理器，用于调用所述存储器中存储的程序指令，按照获得的程序执行上述第一方面所述的方法。
- [0063] 第四方面，本发明实施例提供一种计算机可读非易失性存储介质，包括计算机可读指令，当计算机读取并执行所述计算机可读指令时，使得计算机执行上述第一方面所述的方法。

附图说明

- [0064] 为了更清楚地说明本发明实施例中的技术方案，下面将对实施例描述中所需要使用的附图作简要介绍，显而易见地，下面描述中的附图仅仅是本发明的一些实施例，对于本领域的普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。
- [0065] 图1为本发明实施例提供的一种系统架构的示意图；
- [0066] 图2为本发明实施例提供的一种根因分析的方法的流程示意图；
- [0067] 图3为本发明实施例提供的一种根因分析的方法的示意图；
- [0068] 图4为本发明实施例提供的一种根因分析的方法的流程示意图；
- [0069] 图5为本发明实施例提供的一种根因分析的方法的流程示意图；
- [0070] 图6为本发明实施例提供的一种根因分析的方法的流程示意图；

[0071] 图7为本发明实施例提供的一种根因分析装置的结构示意图。

具体实施方式

[0072] 为了更好的理解上述技术方案,下面将结合说明书附图及具体的实施方式对上述技术方案进行详细的说明,应当理解本发明实施例以及实施例中的具体特征是对本发明技术方案的详细的说明,而不是对本发明技术方案的限定,在不冲突的情况下,本发明实施例以及实施例中的技术特征可以相互结合。

[0073] 首先,对本申请中的部分用语进行解释说明,以便使本领域技术人员理解。

[0074] 业务指标:监控系统中,与业务行为相关的指标,如交易量、成功率、耗时、成功量等指标。

[0075] 业务异常:某一个业务产品的某一个应用场景发生异常,如:微信支付失败。

[0076] 交易流水:一个用于标记一笔交易的编号,可流转在用于处理该交易的多个子系统中。

[0077] 为了使本申请的目的、技术方案和优点更加清楚,下面将结合附图对本申请作进一步地详细描述。

[0078] 图1是本发明实施例所适用的一种系统架构示意图,可以形象的称为业务处理树,本发明是基于树上节点完整的业务指标上报的前提下进行的。其中,业务处理树的构建方法是基于大量的交易流水数据,从而构建出一个业务场景的系统调用树形图。

[0079] 需要说明的是,在这个树形图中,可以由各业务系统自主上报对应的业务指标到树节点上,并为每一个指标绑定相关的属性。比如:对应的外部合作商、关联的机构和银行、对应的后台接口和服务、维度和分类等。

[0080] 进一步的,在完成以上基础工作准备后,可选对树中各节点的业务指标进行曲线异常检测,或者对树种其中一个节点的业务指标进行曲线异常检测,这里可以是根节点,也可以是子节点,本申请对此不做具体限定。此时,可以用曲线异常检测算法来进行检测。本申请为了提高分析的准确性和智能性,优选对树中各节点的业务指标进行曲线异常检测。

[0081] 在一种可能的方案中,比如采用3-sigma及进行检测:先假设一组检测数据只含有随机误差,对其进行计算处理得到标准偏差,按一定概率确定一个区间,认为凡超过这个区间的误差,就不属于随机误差而是粗大误差,对含有该误差的数据予以剔除。

[0082] 本申请实施例中,当发生业务异常时可以发现树中各个子系统之间的指标会有一些异动上的关联关系,比如:各子系统会在异常点附近同时出现上升或者下降。基于这一特点,本申请实施例采用了判断不同曲线异动的相似度的方法进行根因分析。

[0083] 为了更好的解释上述实施例,图2示例性的示出了一种根因分析的方法的流程,该流程可以由根因分析的装置执行。

[0084] 如图2所示,该流程具体包括:

[0085] 步骤201,在监控到第一业务节点处理第一业务产生异常后,获取第一业务节点中发生异常的第一指标曲线。

[0086] 需要说明的是,本申请在异常检测发现异常之后获取发送异常的第一指标曲线。在一种可能设计的方案中,当一个节点上的多条指标曲线中的一条在某一时刻指标不满足预设的条件,此时告警器则会产生一个告警,提示异常。在另一种可能设计的方案中,系统

维护人员根据预设的核心指标进行检测异常,当对应的指标曲线不满足预设的条件时,系统维护人员判断这条指标曲线产生异常。

[0087] 步骤202,确定第一业务对应的第二业务节点。

[0088] 需要说明的是,第二业务节点为处理第一业务的各业务节点中除第一业务节点之外的任一业务节点。

[0089] 步骤203,从各第二业务节点中确定初选异常节点。

[0090] 需要说明的是,初选异常节点中具有与第一指标曲线匹配的第二指标曲线。

[0091] 步骤204,从第一业务节点和初选异常节点中确定出引发异常的根因节点。

[0092] 本申请实施例中,因为入口侧节点的指标一般都比较容易进行产品和业务归属的人工标记,比如:某个指标对应的是微信的访问量,而非入口侧的子节点一般会被多个产品或业务公用,比如:一个存款的子系统可能会有很多不同入口的业务接入。

[0093] 从上述内容可以看出,非入口节点对应的指标很难直接与产品或业务绑定,假使能够找到非入口侧节点指标和业务的对应关系,这样,当某个业务的指标异常时,就可以快速找到该场景所关联的所有指标集合,而不是全量的指标集合(也即包含与该业务没有关联的指标集合),从而达到缩小匹配范围的目的。

[0094] 因此在步骤202中,为了确定第一业务对应的第二业务节点,在步骤201监控到第一业务节点处理第一业务产生异常之前,还要通过以下方式建立指标曲线与业务节点的对照关系:

[0095] 具体的,首先,通过第三业务的业务处理流水,确定处理第三业务的各第三业务节点,同时确定每个第三业务节点的各指标曲线中与第三业务关联的指标曲线。

[0096] 需要说明的是,第三业务为各业务中的任一种。

[0097] 最后,生成第三业务的业务标识、各第三业务节点、各第三业务节点中与第三业务关联的指标曲线的对照关系。

[0098] 从上述内容可以看出,通过第三业务的业务处理流水,可以确定第三业务经过的节点以及第三业务参与了节点上的哪些指标曲线。

[0099] 举个例子,某一个子系统X,在某一时刻有1万条包含交易流水的日志数据,这1万条日志,可根据用户配置的计算规则比如日志过滤规则,计算得到100条指标曲线。其中,如果某一笔日志满足某一个规则,则确定其属于某一个曲线。比如,日志中记载有“success”,说明其与成功率指标相关。

[0100] 基于上述内容,在树的根节点选择一笔交易流水,并考察这笔交易流水在下游的哪些子节点出现过,并参与构成了哪些指标。那么,只要将这一笔流水绑定到一个业务,这样所有的子节点对应的指标就都可以绑定到该业务。通过这种方法得到的指标和业务的对应关系,是1对多的关系,即一个指标可以参与构成多个业务。

[0101] 举一个简单的例子,如图3所示,A交易为微信访问按顺序依次经过了“1”、“4”、“8”三个节点并参与了节点1上成功率的指标,B交易为QQ访问按顺序依次经过了“1”、“5”两个节点,也参与了节点1上成功率的指标,那么节点“1”上成功率对应的指标曲线就会与微信访问、QQ访问这两个业务相关联。

[0102] 进一步的,由于一笔流水只能反映某一个业务的部分情况,也就是说一笔流水不一定能流转所有子节点和所有的指标上,因此以上方案在实施的是时候,是批量进行的。

[0103] 在一种可能的方案设计中,可以搜集入口侧指标对应的一批交易流水编号,优选的,通过离线的方式,统计这些指标在树中各个节点的流转情况,及分析交易流水和节点中指标曲线的构成情况。

[0104] 上述方案,避免了由于相似判断考察的范围过大,导致一些本来不相关的曲线凑巧在相同的时间点出现异动(当样本达到一定量的时候,在某一个时刻总会有一些曲线刚好出现正常的异动)的情况,通过以上的离线分析,相当于是通过染色的方式,找到入口的每一个指标和树中的哪些指标是有关联的,哪些是没有关联的。在做相似判定的时候,则只在有关联中的集合中查找。这样判断相似的集合大大缩小,准确性得到提高。

[0105] 本申请实施例在步骤203中,具体的,针对每个第二业务节点,确定第二业务节点各指标曲线中与所述第一业务关联的指标曲线;若与第一业务关联的指标曲线中存在与第一指标曲线匹配的第二指标曲线,则确定第二业务节点为初选异常节点。

[0106] 进一步的,为了实现上述的判定方法,需要找出树中各个节点中和发生异常的指标曲线指标异动相似的曲线。也就是说,本申请实施例提供了一种方法来判定异常时间段内各个曲线之间的走势是否相似。

[0107] 需要说明的是,本申请实施例并不关注一条曲线整体是否相似或走势相同,而只是关注在指标异常的这个时间点附近,不同曲线是否出现相同波动。

[0108] 具体的,通过如下方式确定是否与第一指标曲线匹配的第二指标曲线,具体流程如图4所示,包括:

[0109] 步骤401,确定第一指标曲线中的第一异常开始时刻、第一异常峰值时刻及异常时段中各时刻的指标波动值的第一排序;

[0110] 步骤402,确定第二指标曲线中的第二异常开始时刻、第二异常峰值时刻及异常时段中各时刻的指标波动值的第二排序;

[0111] 步骤403,根据第一异常开始时刻与第二异常开始时刻的第一匹配度、第一异常峰值时刻与第二异常峰值时刻的第二匹配度及第一排序与第二排序的第三匹配度,确定第一指标曲线与第二指标曲线之间的匹配度是否大于设定阈值。

[0112] 本申请实施例中,选取了曲线的三个特征,具体包括异常开始时刻、异常峰值时刻及异常时段中各时刻的指标波动值的排序,并利用这三个特征来判断相似性。需要说明的是,还可选取异常结束时刻、异常持续时间、异常时间段内指标最大值与最小值的差值等其它特征来判断相似性,本申请对此不做具体限定,下面分别描述一下这三个特征的获取方法:

[0113] 异常开始时刻:采用曲线异常检测算法比如:3-Sigma,找出异常的开始时间。

[0114] 异常峰值时刻:找出时间段内交易量、成功率的最低值,或者耗时的最高值对应的时间点。

[0115] 指标波动值的排序:针对曲线中的每一个点,找出该点的前N个点对应的平均值,并用该值减去之前的平均值,得到该点的差分值即波动值。找出曲线异常时间点前X个小时所有点的波动值,并对波动值进行排序,得到当前曲线对应的波动值的排序。

[0116] 针对上述指标波动值的排序举个例子,比如找到时刻为18:20的点,获取该点之前20分钟所包含的点的平均值,比如每间隔一分钟采集一个点,此时得到20个点的指标值,取平均值,再将平均值从大到校进行排序。

[0117] 进一步的,在完成了以上三个特征的采集后,可以通过某条曲线的特征和发生异常的指标曲线特征的匹配程度来打分。在一个可能的方案设计中,开始时间完全相同计3分,每相差1分钟减1分。峰值时刻完全相同计3分,每相差1分钟减1分。波动值的排序完全相同计3分,每相差1减1分。如果某一个类值为负分,则直接排除。

[0118] 从上述内容可以看出,通过以上的打分加权求和,可以得到某一条曲线和发生异常的指标曲线的相似打分。当高于一定的阈值时,将其确定为初选异常节点。

[0119] 从上述内容可以看出,从第一业务对应的第二业务节点各指标曲线中确定出与前述第一业务关联的指标曲线,将与所述第一业务关联的指标曲线进行匹配;由于实际业务中,一个业务节点会涉及多种业务,而且一个业务节点上会基于各种需求设置很多的指标曲线;本方案中通过缩小业务节点的范围,缩小指标曲线的范围,可以提高了根因分析的效率;同时,由于排除了不相关的曲线,从而避免了某些曲线由于波动导致了与第一指标曲线的匹配度提升,从而也提高了根因分析的准确度。

[0120] 然而通过上述曲线相似的判断方法仍会发现一些误判的情况。也就是说根因分析的结果经过核查后,发现并不是真正的根因。出现误判的原因可能是由于一些被判断为相似异动的曲线,从逻辑层面来考察,也发现数据上的不合理情况,比如:入口耗时增加1000ms,核心侧只增加了20ms,即核心侧20ms的增长导致了入口侧1000ms的增长,这显然是不合理的。基于此,本申请实施例在步骤204中,具体通过如图5所示的方法流程排除初选异常节点中的一部分节点。

[0121] 具体流程如下:

[0122] 步骤501,针对发生异常的同—指标,确定所第一业务节点在指标的第一变动量,及初选异常节点在指标的第二变动量。

[0123] 步骤502,若第一变动量与第二变动量之间的关系不满足设定条件,则排除初选异常节点。

[0124] 步骤503,从排除操作后的初选异常节点和第一业务节点中确定出引发异常的根因节点。

[0125] 上述方案,从判断数据相关的合理性的较低,来进一步提高根因分析的准确性。

[0126] 举个例子,先通过上述内容,分析出所有异动相似的曲线。当出现入口侧成功率降低时,可以找到计算当前成功率用到的成功量和总量即成功率=成功量/总量,并可以进一步得到失败量即失败量=总量-成功量,通过该方法,可以找到所有相似曲线对应的失败量。如果某一个节点失败量,相较于发生异常的节点的失败量太低,则表明,不可能是因为该节点的失败,导致发生异常的节点的大量失败,该相似节点被作为根因的可能性被排除;类似的耗时升高时,也可以做类似的处理,本申请在此不多做赘述。

[0127] 比如,异常发生在入口节点,入口节点的交易量下降了1000,而初选异常节点只下降了10,说明该初选异常节点不是主要原因,则排除。

[0128] 上述方法,可以排除掉一些原本被判定相似,但不合理的曲线,提高判定的准确性。

[0129] 进一步的,本申请实施例在确定了初选异常节点之后,针对第一指标曲线对应的指标,从初选异常节点和第一业务节点中确定出在处理第一业务时在指标上的影响方,将影响方确定为引发异常的根因节点。

[0130] 具体来说,如果是入口侧交易量出现下降,及后面的部分子系统的交易量都下降,则可以判定是因为入口侧流量减小,导致的后端交易量指标减小,根因是入口节点。

[0131] 如果是入口侧出现了成功率突降或者耗时突增的异常,及后面的部分子系统也出现相同业务指标的异常,则可以判断是因为后端的成功率降低或耗时增加,导致了入口侧的成功率降低或耗时增加,根因是最后一个异常的树节点,也就是说,后面的节点出现的成功率降低或耗时增高必定会反馈到前面的节点中。

[0132] 上述方案,根据初选异常节点与发生异常节点之间的依赖关系确定最终的根因节点,减少了误判的可能,提高了根因分析的准确性。

[0133] 进一步的,通过以上方案的RCA判断后,可以得到根因节点及根因的指标,然而这些信息,对除了系统开发者以外的人来说,可读性不好。本申请实施例通过以下方法,来进一步提高根因分析结论的可读性,具体流程如图6所示。

[0134] 步骤601,确定所第一指标曲线对应的指标所关联的第一属性及第二指标曲线对应的指标所关联的第二属性。

[0135] 步骤602,展示第一属性和第二属性中的相同属性。

[0136] 具体的,找到所有相似异动的指标及指标所关联的属性,找出共同点。比如,指标均关联到某个外部合作方、某一家银行等。通过这些信息的统计和输出,可以将异常圈定在一个更小的范围或分类中。

[0137] 进一步的,本申请实施例中查找关联属性的方法,在一种可能的实现方式中,通过用户上传指标时,在上报数据中绑定的相关属性,在另一种可能的实现方式中,通过在管理控制台在完成上报后,编辑相关属性。除此之外,通过指标的名称,也可能会模糊匹配到银行或机构的名称。比如,一个指标名称可能为:建行接口调用成功率,通过将该指标与一个银行或银行缩写字典绑定,则可以在名称中提取出相关银行为:建行。

[0138] 在一种可能的实现方式中,针对系统成功率降低的异常,可以事先将所有指标对应的失败交易日志的相关信息入库,形成错误原因数据信息库,当成功率降低时,就可以在该信息库中找到所有指标中引发成功率降低的相关日志,并从日志中提取相关信息,举例来说,日志中涉及的中文,特定的分类比如连接超时、数据库访问异常等。通过相关日志的统计和输出,也可以更好的让用户明白问题的原因。

[0139] 上述方案,通过找到所有相似异动的指标及指标所关联的属性并进行统计和输出,可以将异常圈定在一个更小的范围或分类中。提高了根因分析结果的可读性。从宏观的角度来看待问题,能够更准确、更全面的发现问题的根源,能够更好的找到问题的分类或指向,通过这种分类或指向,再结合微观分析,能更好的得出根因分析结论。

[0140] 基于相同的技术构思,图7示例性的示出了本发明实施例提供一种根因分析的装置的结构,该装置可以执行根因分析的方法的流程。

[0141] 如图7所示,该装置可以包括:

[0142] 获取模块701,用于在监控到第一业务节点处理第一业务产生异常后,获取所述第一业务节点中发生异常的第一指标曲线;

[0143] 处理模块702,用于确定所述第一业务对应的第二业务节点,所述第二业务节点为处理所述第一业务的各业务节点中除所述第一业务节点之外的任一业务节点;

[0144] 所述处理模块702,还用于从各第二业务节点中确定初选异常节点,其中,所述初

选异常节点中具有与所述第一指标曲线匹配的第二指标曲线；

[0145] 所述处理模块702,还用于从所述第一业务节点和所述初选异常节点中确定出引发异常的根因节点。

[0146] 可选的,所述处理模块702具体用于:

[0147] 针对每个第二业务节点,确定所述第二业务节点各指标曲线中与所述第一业务关联的指标曲线;若与所述第一业务关联的指标曲线中存在与所述第一指标曲线匹配的第二指标曲线,则确定所述第二业务节点为初选异常节点。

[0148] 可选的,所述处理模块702还用于:

[0149] 在所述在监控到第一业务节点处理第一业务产生异常之前,通过第三业务的业务处理流水,确定处理所述第三业务的各第三业务节点,并确定每个第三业务节点各指标曲线中与所述第三业务关联的指标曲线;所述第三业务为各业务中的任一种;

[0150] 生成所述第三业务的业务标识、各第三业务节点、各第三业务节点中与所述第三业务关联的指标曲线的对照关系。

[0151] 可选的,所述处理模块702具体用于:

[0152] 确定所述第一指标曲线中的第一异常开始时刻、第一异常峰值时刻及异常时段中各时刻的指标波动值的第一排序;

[0153] 确定所述第二指标曲线中的第二异常开始时刻、第二异常峰值时刻及异常时段中各时刻的指标波动值的第二排序;

[0154] 根据所述第一异常开始时刻与所述第二异常开始时刻的第一匹配度、所述第一异常峰值时刻与所述第二异常峰值时刻的第二匹配度及所述第一排序与所述第二排序的第三匹配度,确定所述第一指标曲线与所述第二指标曲线之间的匹配度是否大于设定阈值。

[0155] 可选的,所述处理模块702具体用于:

[0156] 针对发生异常的第一指标,确定所述第一业务节点在所述第一指标的第一变动量,及所述初选异常节点在所述第二指标的第二变动量;

[0157] 若所述第一变动量与所述第二变动量之间的关系不满足设定条件,则排除所述初选异常节点;

[0158] 从排除操作后的初选异常节点和所述第一业务节点中确定出引发异常的根因节点。

[0159] 可选的,所述处理模块702具体用于:

[0160] 针对所述第一指标曲线对应的指标,从所述初选异常节点和所述第一业务节点中确定出在处理所述第一业务时在所述指标上的影响方,将影响方确定为引发异常的根因节点。

[0161] 可选的,其特征在于,所述处理模块702还用于:

[0162] 确定所述第一指标曲线对应的指标所关联的第一属性及所述第二指标曲线对应的指标所关联的第二属性;

[0163] 展示所述第一属性和所述第二属性中的相同属性。基于相同的技术构思,本发明实施例还提供了一种计算被控设备,包括:

[0164] 存储器,用于存储程序指令;

[0165] 处理器,用于调用所述存储器中存储的程序指令,按照获得的程序执行上述筛选

机器学习模型的方法。

[0166] 基于相同的技术构思,本发明实施例还提供了一种计算机可读非易失性存储介质,包括计算机可读指令,当计算机读取并执行所述计算机可读指令时,使得计算机执行上述筛选机器学习模型的方法。

[0167] 最后应说明的是:本领域内的技术人员应明白,本发明的实施例可提供为方法、系统、或计算机程序产品。因此,本发明可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且,本发明可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、光学存储器等)上实施的计算机程序产品的形式。

[0168] 本发明是参照根据本发明的方法、被控设备(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被控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被控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和/或方框图一个方框或多个方框中指定的功能的装置。

[0169] 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被控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造品,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。

[0170] 显然,本领域的技术人员可以对本发明进行各种改动和变型而不脱离本发明的范围。这样,倘若本发明的这些修改和变型属于本发明权利要求及其等同技术的范围之内,则本发明也意图包含这些改动和变型在内。

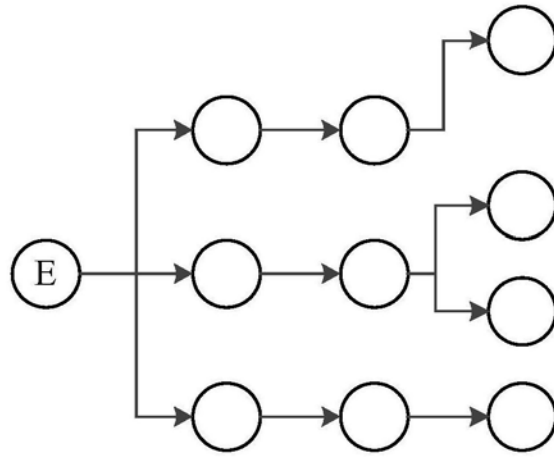


图1

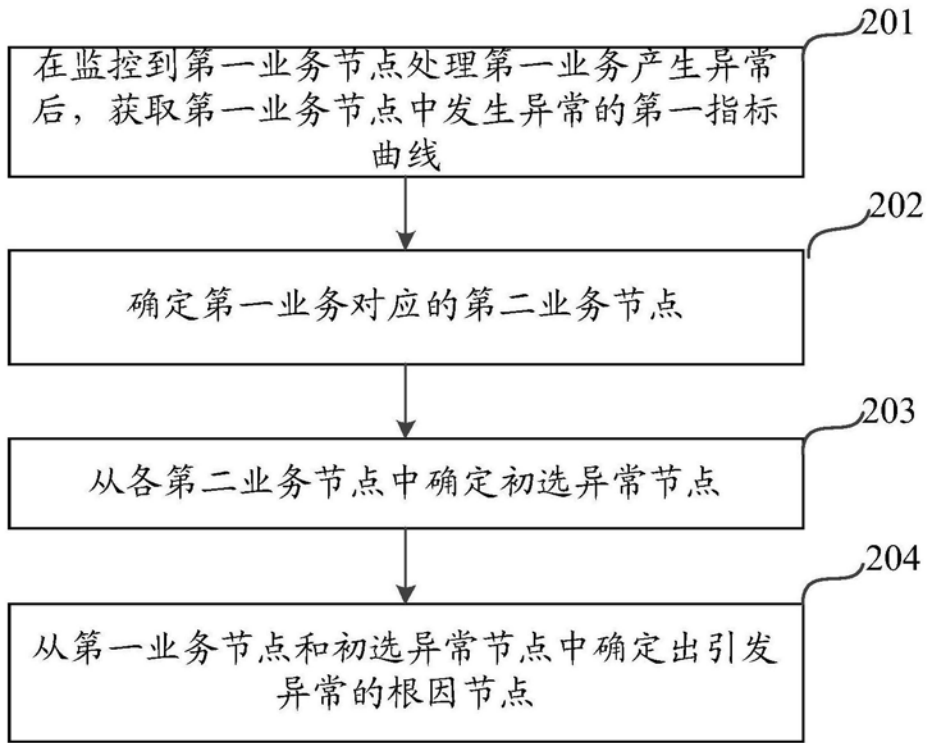


图2

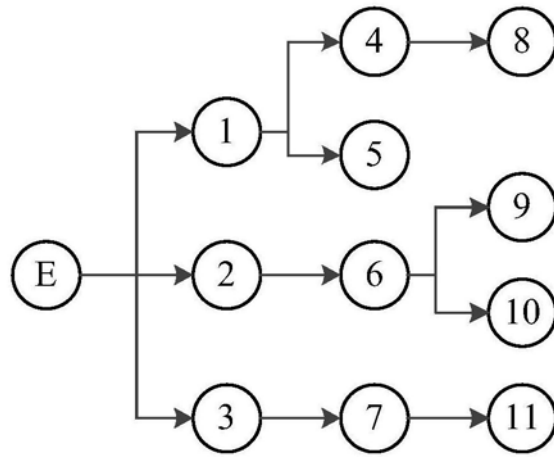


图3

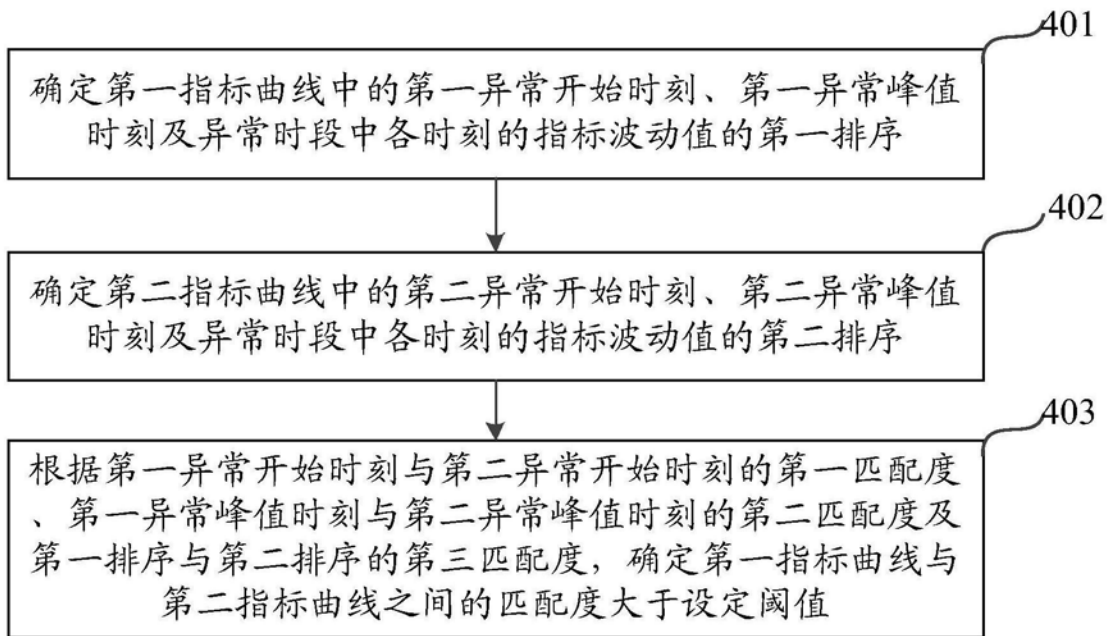


图4

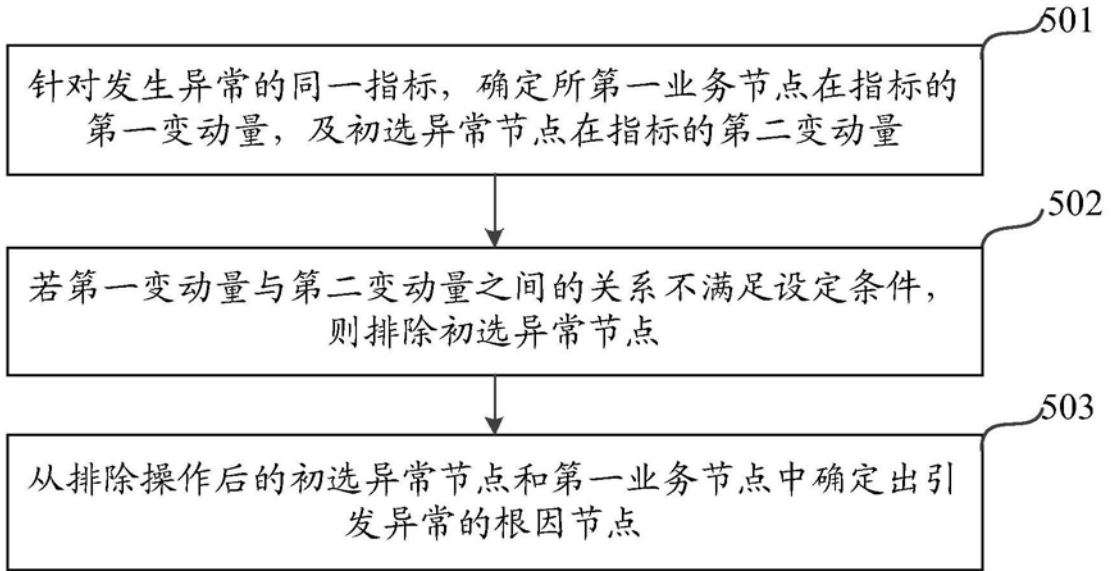


图5

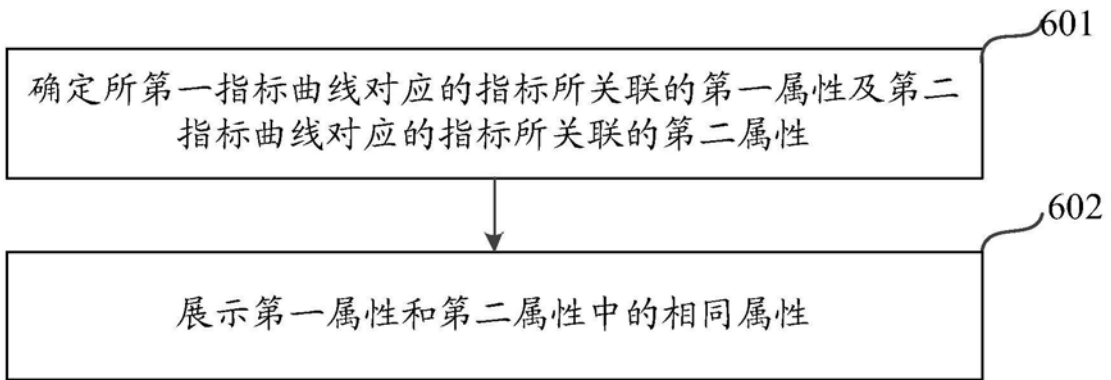


图6

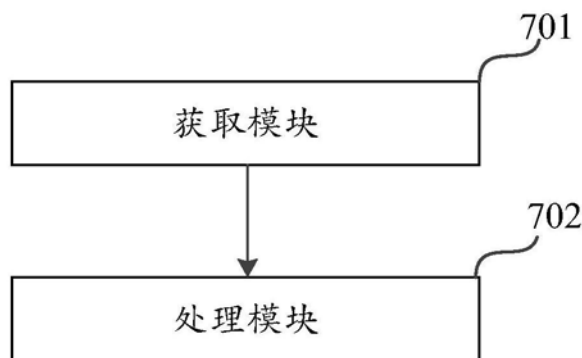


图7